

Pred. No.:	2,16e-177	Length:	514
Score:	2405.00	Matches:	460
Percent Similarity:	92.66%	Conservative:	20
Best Local Similarity:	88.80%	Mismatches:	34
Query Match:	72.03%	Indels:	4
DB:	11	Gaps:	1

US-09-806-194-1 (1-1804) x Q8C5E9 (1-514)

Qy	1	ATGGGCGCACTGCGCCGGCGCTGCTGCTGCTGCTGCCAGTGGCTCTGTCCGCGC	60
Dd	1	MetGlyAlaLeuLeuArgAlaLeuLeuLeuProValLeuAlaGlnTrpLeuLeuSerAla	20
Qy	61	GGCCGGAGCTGGCCCGCCGCCCTTCACGTGCCCCICGGGGTGCCGCGGCACGAAC	120
Dd	21	ValProAlaLeuAlaProPheThrLeuProLeuGlnValAlaArgAlaThrAsn	40
Qy	121	CQGTRAGTTGCCACCACCGGGACCCCGGACCCCTCCGAGCCGCCGCGAGCGCTTG	180
Dd	41	HisArgAlaSerAlaValProGlyLeuGlyThrProGlyLeuProArgAlaAspGlyLeu	60
Qy	181	GGGCTCGCCCTGGAGCCTGGCTGGCTGCCGCCGCCGCGCGCAAATCTTGTGGCCATG	240
Dd	61	AlaLeuAlaLeuGluProValArgAlaThr-----AlasnPheLeuAlaMet	76
Qy	241	GTAGACAACCTCGAGGGGACTCTGGCGCGGCTACTACCTGGAGATGCTGATCGGGACC	300
Dd	77	ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrrLeuGluMetLeuIleGlyThr	96
Qy	301	CCC CGCAGAAGCTACAGATTCTGTTGCACACTCGAAGCAGTAATTTGCCGTGGCAGA	360
Dd	97	ProProGlnLysValGlnLeuValAspThrGlySerSerAsnPheAlaValAlaGly	116
Qy	361	ACCCGCGCATCTACATAGACACGCTACTTTGCACACAGAGAGGTCTAGCACATACCCGCTCC	420
Dd	117	AlaProHisSerTyrIleAspThrTyrPheAspSerGluSerSerSerThrTyrHisSer	136
Qy	421	NAGGGCTTGAGCTCACAGTCAAGTACACACAGAAGCTGGACGGCTTCGTGGSGAA	480
Dd	137	LysGlyPheAspValThrValLysTyrThrGlnGlySerrTrpThrGlyPheValGlyGlu	156
Qy	481	GACCTCGTCACCATCCCCAAGGTTCAATACTTCTTTCTTGTCAACATTCACATATT	540
Dd	157	AspLeuValThrIleProLysGlyPheAsnSerSerPheLeuValAsnIleAlaThrIle	176
Qy	541	TTTGATCATAGAAATTTCTTTTGCTGGGATTAAATGGAATGGAATACTTGGCTTAGCT	600
Dd	177	PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla	196
Qy	601	TATGCCACACTTGCACAAGCATCAAGTTCTCTGGAGACCTTCITCGACTCCCTGGTGACA	660
Dd	197	TyrAlaLeuAlaLysProSerSerLeuGluThrPhePheAspSerLeuValAla	216
Qy	661	CAAGCAAAATCCCCAACGTTTTCTCCATGCAGATGTGTGAGCGCGCTTGCCTGTGT	720
Dd	217	GlnAlaLysIleProAspIlePheSerMetGlnMetCysGlyAlaGlyLeuProValAla	236
Qy	721	GNATCTGGACCAACGAGGTAGCTTGTCTGGTGGAAATTAACCAAGTTGTGTATAA	780
Dd	237	GlySerGlyThrAsnGlyLysLeuValLeuGlyGlyIleGluProSerLeuTrpLys	256
Qy	781	GGAGACATCTGGTATACCCCTATTAAAGGAAGTGGTACTACCAGATAGAAAATCTGAAA	840
Dd	257	GlyAspIleerPrThrProLysGluLutprTyrGlnIleGluIleLeuLys	276
Qy	841	TTTGAATTTGGAGGCCAAGCCTTAATCTGACTGCAGAGAGTATAACGCAGACAAGGCC	900
Dd	277	LeuGluIleGlyGlyGlnAsnLeuAsnLeuAspcysargGluTrpAsnAlaAspLysAla	296
Qy	901	ATCGTGAGCAGTGGCACACGCTGCTGGCTCGCCCGAAGGTGTTTGTATGCGGTGGT	960
Dd	297	IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValal	316

QY	961	GAAGCTGTGGCCCGCCGATCTCTGATTCACAGAAATCTCTGATGTTTCTGGACTGGGTCC	1020
DB	317	GUUUAValAlaArgThrSerLeuLeuProGluPheSerAspGlyPheTrpThrGlyAla	336
QY	1021	CAGCTGGCTGCTGGAGGAATTCGGAACACCTTGGTCTTACTTCCCTAAATCTCCATC	1080
DB	337	GluLeuAlaCysTrpThrAsnSerGluThrProTrpAlaTrpPheProLysIleSerIle	356
QY	1081	TACCTGAGAGATCAGAACTCCACAGGTCAATCCGTATACAACTCCGTCCCTCAGCTTTAC	1140
DB	357	TyrLeuArgAspGluAsnAlaSerArgSerPheArgIleThrIleLeuProGlnLeuTrp	376
QY	1141	ATTGAGCCCATGATGGGGCCGGCTGAATATGAATTAACGATTACCGATTTCGCCCA	1200
DB	377	IleGlnProMetMetGlyAlaGlyPheAsnTyrGluCysTyrArgPheGlyIleSerSer	396
QY	1201	TGCACAAATCCGCTGGTGATCGGTGCCACGGTGATGGAGGGTTCTACGTATCTTCGAC	1260
DB	397	SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValValPheAsp	416
QY	1261	AGAGCCCAGAAAGGGTGGCTTCGCGAGCAGCCCTGTCAGAGAAATTCAGAGTGCTGCA	1320
DB	417	ArgAlaGlnArgValGlyPheAlaValSerProCysAlaGluIleGluGlyThrThr	436
QY	1321	GTGCTGAAATTTCCGGCGCTTTCTCAACAGAGGATGATGACCAAGCAACTGTGCCCCGT	1380
DB	437	ValSerGluIleSerGlyProPheSerThrGluAspIleAlaSerAsnCysValProAla	456
QY	1381	CAGTCTTTGACGAGCCCATTTGTGGATTGTGCTCTATCGGCTCAGAGCGGTCTGTGGA	1440
DB	457	GluAlaLeuAsnGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly	476
QY	1441	GCCATCCCTCTGCTTAATCGTCTCTGCTGCTGCGCGTTCGGGTCTCAGCGTCGCCCC	1500
DB	477	AlaIleLeuLeuValLeuIleLeuLeuLeuLeuLeuProLeuHisCysArgHisAlaPro	496
QY	1501	CGTGACCTCAGGTGTCATATGATGAGTCTCTCTGTGTCAGACATCGCTGGAAA	1554
DB	497	ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys	514

RESULT 2

Q8C793 PRELIMINARY; PRT; 514 AA.

ID	Q8C793	PRELIMINARY;	PRT;	514 AA.
AC	Q8C793;			
DT	01-MAR-2003 (TrEMBLrel, 23, Created)			
DI	01-MAR-2003 (TrEMBLrel, 23, Last sequence update)			
DE	01-MAR-2003 (TrEMBLrel, 23, Last annotation update)			
DE	Beta-site APP-cleaving enzyme 2.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-C57BL/6J; TISSUE=Heart;			
RX	MEDLINE=22354683; PubMed=12456851;			
RA	The FANTOM Consortium,			
RA	the RIKEN Genome Exploration Research Group Phase I & II Team;			
RT	"Analysis of the mouse transcriptome based on functional annotation of			
RT	60,770 full-length cDNAs";			
RJ	Nature 420:563-573(2002).			
DR	ENBL; AK052309; BAC34931.1; -.			
SQ	SEQUENCE 514 AA; 55871 MW; 8BF45E07B0990225 CRC64;			

Alignment Scores:

Pred. No.:	6,29e-177	Length:	514
Score:	2399.00	Matches:	459
Percent Similarity:	92.47%	Conservative:	20
Best Local Similarity:	88.61%	Mismatches:	35
Query Match:	71.85%	Indels:	4
DB:	11	Gaps:	1

US-09-806-194-1 (1-1804) x Q8C793 (1-514)

```
QY 1 ATGGCGCACTGGCCGGCGGCTGCTGCTCCCTCTGCTGGCCAGTGGCTCTGGCGCC 60
Db 1 MetGlyAlaLeuLeuArgAlaLeuLeuLeuProValLeuAlaGlnTrpLeuLeuSerAla 20
QY 61 GCCCGAGAGTGGCCCGCCCGCTTCAGCGTCCCTCCGCGTGGCGCGCCGACCAAC 120
Db 21 ValProAlaLeuAlaProAlaProPheThrLeuProLeuGlnValAlaArgAlaThrAsn 40
QY 121 CGCGTAGTTCGGCCCGCCCGGAGCCCGGAGCCCTCGCCGAGCCCGCCAGCGGCTG 180
Db 41 HisArgAlaSerAlaValProGlyLeuGlyThrProGluLeuProArgAlaAspGlyLeu 60
QY 181 CGCGTCCCGCTGGAGCGTGGCGGCTCCCGCGGCGCGCGCCCAACTTCCTGGSCCATG 240
Db 61 AlaLeuAlaLeuGluProValArgAlaThr-----AlaAsnPheLeuAlaMet 76
QY 241 GTAGACAACCTGGAGGGGACTCTGGCGGCTACTACTGAGAGTGTGATCGGGACC 300
Db 77 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuThrGlyThr 56
QY 301 CCCCCGAGAGCTACAGATTCTGTTGACACTGGAAGCAGTAACTTTCGCGTGGCAGGA 360
Db 97 ProProGlnLysValGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 116
QY 361 ACCCGCACTCTACATAGACAGTACTTTGACACAGAGAGTCTACACATACCGCTCC 420
Db 117 AlaProHisSerTyrIleAspThrTyrPheAspSerGluSerSerSerThrTyrHisSer 136
QY 421 AAGGSGTTGACGTCACAGTCAAGTACACACAGGAGCTGGACGGCTTCGTTGGGAA 480
Db 137 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 156
QY 481 GACCTGCTCACTCCCAAGGCTCAATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 540
Db 157 AspLeuValThrIleProLysGlyPheAsnSerSerPheLeuValAsnIleAlaThrIle 176
QY 541 TTTCGAATCAGAGATTTCTTTTTCCTGGGATTAATGATGAATAGTACTTGGCCCTAGCT 600
Db 177 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 196
QY 601 TATCCCACTGTGCAAGCCTCAAGTCTCTGAGACCTTCTTCGACCTCCCTGGTGACA 660
Db 197 TyrAlaAlaLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValAla 216
QY 661 CAAGCAACATCCCAACGTTTCTCCATGCAGATGGTGGAGCGCGCTGGCCCTTGCT 720
Db 217 GlnAlaLysIleProAspIlePheSerMetGlnMetCysGlyAlaGlyLeuProValAla 236
QY 721 GGATCTGGGACCAACGAGGTAGTCTTGTGGTGGAAITCAACCAAGTTTGTATATA 780
Db 237 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 256
QY 781 GGAGACATCTGGTATACCCCTATTAGGAAGATGGTACTACCAGATAGAAATTCGTAA 840
Db 257 GlyAspIleTrpTyrThrProIleLysLysLysLysLysLysLysLysLysLysLys 276
QY 841 TTGGAATTCAGGCGCAACCCCTAATCTGCAGTGGCAGAGATATACGACAGACAGCC 900
Db 277 LeuGluIleGlyGlyGlnAsnLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 296
QY 901 ATCTGGACAGTGCACCGCTCTCGGCTCGCCCGAGAGGTGTTTGAATGGCGGTGGTG 960
Db 297 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 316
QY 961 GAAGCTGGCGCCCGCATCTCTGATTCAGAAATCTCTGATGTTTCTGGATGGGTCC 1020
Db 317 GluAlaValAlaArgThrSerLeuIleProGluPheSerAspGlyPheTrpThrGlyAla 336
QY 1021 CAGCTGGCGCTCGGAGCAATTCGGAACACCTTGGTCTTACTTCCCTAAATCTCCATC 1080
Db 337 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpAlaTyrPheProLysIleSerIle 356
```

```
QY 1081 TACCTGAGAGTGAAGAACTCCAGCAGGTCTATTCGGTATACAACTCTGCTCAGCTTAC 1140
Db 357 TyrLeuArgaspGluAsnAlaSerArgSerPheArgThrThrIleLeuProGlnLeuTyr 376
QY 1141 ATTCAGCCCCATGATGGGCGCGCTGAATATGAATTTACCGATTCGGCATTTCCCA 1200
Db 377 IleGlnProMetMetGlyAlaGlyPheAsnTyrGluCysTyrArgPheGlyIleSerSer 396
QY 1201 TCCCAAAATCGCTGATCGGTGGCCAGCTGATGGAGGCTTCTACGTCATCTTCGAC 1260
Db 397 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValValPheAsp 416
QY 1261 AGACCCCAAGAGGTGGCTTCGACAGCCCGCTGTCAGAAAATTCAGGTGCTGCA 1320
Db 417 ArgAlaGlnArgValGlyPheAlaValSerProCysAlaGluIleGluGlyThrThr 436
QY 1321 GTGCTGAAATTTCCGGCGCTTTCTCAACAGAGATCTAGCCAGCACTGTGTCGCCGCT 1380
Db 437 ValSerGluIleSerGlyProPheSerThrGluAspIleAlaSerAsnCysValProAla 456
QY 1381 CAGTCTTTCAGCGAGCCCAITTTGTGATTGTGTCTCTATCGCTCATGACGCTCTGTGA 1440
Db 457 GlnAlaLeuAsnGluProIleLeuTyrIleValSerTyrAlaLeuMetSerValCysGly 476
QY 1441 GCCATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1500
Db 477 AlaIleLeuLeuValLeuIleLeuLeuLeuLeuLeuProLeuHisCysArgHisAlaPro 496
QY 1501 CGTCACTGAGTGGTCAATGATGAGTCTCTCTGCTCAGACATCGCTCGAAA 1554
Db 497 ArgaspProGluValValAsnAspSerSerLeuValArgHisArgTrpLys 514

RESULT 3
Q5JL18 PRELIMINARY: PRT; 514 AA.
ID Q9JL18; AC Q9JL18;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DI 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DI 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Aspartyl protease 1.
GN BACE2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Choi D.K., Sugano S., Sakaki Y.;
RT "Molecular characterization of the mouse Aspl gene, a homolog of the
human Aspl (Down Syndrome Region aspartyl protease).";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
DR EMBL: AF216310; AAF36599.1; -.
DR HSSP: P00797; 2REN.
DR MEROPS: A01.041; -.
DR MGD: MGI:1860440; Bace2.
DR InterPro: IPR001461; AspproteaseA1.
DR InterPro: IPR001969; Aspprotease_site.
DR Pfam: PF00026; asp; 1.
DR PRINTS: PR00792; PEPsin.
DR PROSITE: PS00141; ASP_PROTEASE; 2.
KW Aspartyl protease; Hydrolase; Protease.
SQ SEQUENCE 514 AA; 55799 MW; A70725F2C1DF5B47 CRC64;

Alignment Scores:
Pred. No.: 1,28e-176 Length: 514
Score: 2395.00 Matches: 459
Percent Similarity: 92.47% Conservative: 20
Best Local Similarity: 88.61% Mismatches: 35
Query Match: 71.73% Indels: 4
DB: 11 Gaps: 1

US-09-806-194-1 (1-1804) x Q9JL18 (1-514)
```


Db 1 MetGlyAlaLeuAlaArgAlaLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20
 QY 61 GCCCGGAGCTGGCCCGCCGCTTACAGCTGCTCCGCTCGGCTGGCCGCGCAGAAC 120
 Db 21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40
 QY 121 CGCTAGTGTGGCCACCCCGGAGCCCGGAGCCCGCTCCGAGCGCCACGCGCGCTTG 180
 Db 41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60
 QY 181 GCGTCGCGCTGGAGCTGCGCTGCTCCGCGCGCGCGCCGCGCAGCTTCTTGGCAG 240
 Db 61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80
 QY 241 GTAGACAACCTGCGAGGGGACTCTGGCGGGGCTACTACCTGGAGATGCTGATCGGAGC 300
 Db 81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100
 QY 301 CCCCGGAGAGCTACAGATTCCTGACACTGGAGAGAGTACTTTCGCTGGCGAGCA 360
 Db 101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120
 QY 361 ACCCGGACTCCTACATAGACAGCTACTTTCACACAGAGAGTCTAGCACATACCGCTCC 420
 Db 121 ThrProHisSerIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140
 QY 421 AAGCGCTTTGACGTCACAGTGAAGTACACAGAGAGCTGGAGGGCTTCGTTGGGGAA 480
 Db 141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTyrThrGlyPheValGlyGlu 160
 QY 481 GACCTCGTCACATCCCAAGGCTTCAATACTTCTTCTGTCACATTCGCACTATT 540
 Db 161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180
 QY 541 TTGCAATCAGAGAATTTCTTTTCCCTGGGATTAATGGAATGCAATACCTTGCCTAGC 600
 Db 181 PheGluSerGluAsnPhePheLeuProGlyIleLysTyrAsnGlyIleLeuGlyLeuAla 200
 QY 601 TATGCCACACTGCGACGCCATCAAGTCTCTGAGAGCTTCTTCGACTCCCTGGTGACA 660
 Db 201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220
 QY 661 CAAGCAACATCCCAAGCTTTTCTCCATGCAGATGTGTGGAGCGGCTTGCCCGTGTCT 720
 Db 221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240
 QY 721 GATCTGGGACCAAGGAGGTAGTCTTGTCTTGGTGGATTAACCAAGCTTTGTATAA 780
 Db 241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260
 QY 781 GGAGACATCTGTTATACCCCTATTAAAGCAGAGTGGTACTACAGATAGAAATTCGTAA 840
 Db 261 GlyAspIleTyrThrProIleLysGluGluTyrTyrGlnIleGluIleLeuLys 280
 QY 841 TTGGAAATGGAGCCCAAGCTTAATCTGAGTGCAGAGATATAACCGCAGACAGGCC 900
 Db 281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300
 QY 901 ATCTGGGACGTGGCAGCCGCTCGCTCGCTCGCCAGAGGCTGTGTGATGGCTGGT 960
 Db 301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValAla 320
 QY 961 GAAGCTGTGGCCCGGCTCTCTGATTCAGATTCCTGATGCTTCTGATGGTTCCTGGAGTGC 1020
 Db 321 GluAlaValAlaArgAlaSerLeu----- 328
 QY 1021 CAGCTGGCGTCTGGACCAATTCGGAACACCTTGGTCTTACCTCCCTAAATCTCCAIC 1080
 Db 328 ----- 328
 QY 1081 TACCTGAGAGATGAGAACTCCAGCAGGCTATTCGCTATCACAATCTGCTCAGCTTAC 1140
 Db ----- 1140

Db 329 -----LeuTyr 330
 QY 1141 ATTGAGCCCATGATGGCGCGCTGAATTATCAATGTTTACCATTTCCGATTTCCCA 1200
 Db 331 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 350
 QY 1201 TCCAAATGCGTGTGATCGGTGCGCACCGGTGATGAGGGCTTCTACGTCATCTTCGAC 1260
 Db 351 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 370
 QY 1261 AGACCCAGAGAGGTGGCTTCGACGAGCCCTGTGTCAGAAATTCAGGTGCTGCA 1320
 Db 371 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAla 390
 QY 1321 GTGTCTGAAATTTCCGGCGCTTCTCAACAGAGAGATAGCCAGCAACTGTGTCCCGCT 1380
 Db 391 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 410
 QY 1381 CAGTCTTTGACGAGCCCATTTTGTGATTTGTCTCTATGCTGCTCATGACGCTCTGGA 1440
 Db 411 GlnSerLeuSerGluProIleLeuTyrIleValSerTyrAlaLeuMetSerValCysGly 430
 QY 1441 GCCATCTCTCTTCTTAATCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
 Db 431 AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 450
 QY 1501 CGTGACCTGAGCTGCTCAATGATGATGATGATGATGATGATGATGATGATGATGATA 1554
 Db 451 ArgAspProGluValValAlaAsnAspGluSerLeuValArgHisArgTyrLys 468

RESULT 5
 Q9H2V8 PRELIMINARY; PRT: 439 AA.
 ID Q9H2V8 AC Q9H2V8
 DT 01-MAR-2001 (TRENBLrel. 16, Created)
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
 DE CDAL3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Phocchromocytoma;
 RA Li Y., Huang Q., Peng, Y., Song H., Yu Y., Xu S., Ren S., Chen Z.,
 Han Z.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
 DR EMBL; AF212252; AAG41783.1; -.
 DR HSSP; P00797; 2REN.
 DR InterPro; IPR001461; AspproteaseA1.
 DR InterPro; IPR001969; Aspprotease_site.
 DR Pfam; PF00026; asp; 1.
 DR PRINTS; PS00792; PPSIN.
 DR PROSITE; PS00141; ASP_PROTEASE; 2.
 KW Aspartyl protease; Hydrolase; Protease.
 SQ SEQUENCE 439 AA; 48275 MW; 02EC0E050F11602 CRC64;

Alignment Scores:
 Pred. No.: 9,45e-169 Length: 439
 Score: 2293.00 Matches: 439
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 68.67% Indels: 0
 DB: 4 Gaps: 0

US-09-806-194-1 (1-1804) x Q9H2V8 (1-439)

QY 238 ATGTAGACAACTGACGGGGGACTCTGGCGGGCTACTACCTGGAGATGCTGATCGG 297
 Db 1 MetValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGly 20

```
QY 298 ACCGCCCGCCAGACGTACAGATTCGCTGTGACACTGGAAGCAGTAACATTGCCGTGGCA 357
Db 21 ThrProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAla 40
QY 358 GGAACCCCGCCTCTACATACACACGTACTTTGACACAGAGAGCTCTAGCACATACCGC 417
Db 41 GlyThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerThrTyrArg 60
QY 418 TCCAAGGGTTTCAGCTCAGCTCAGAGTACACACAGGAGCGTGGCTTCGTTGGG 477
Db 61 SerLysGlyPheAspValThrValLysTyrThrGlnGlySerThrGlyPheValGly 80
QY 478 GAAGACTCGTCACCATCCCAAGGCTTCAATACTTCTTTCTGTCAACATTGCCACT 537
Db 81 GluAspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThr 100
QY 538 ATTTTGAATCAGAAATTTCTTTTGGCTGGGATTAAATGAAATGGAAATCTGGCCCTA 597
Db 101 IlePheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeu 120
QY 598 GCTTATGCCACACTTGCACAGCCATCAAGTCTCTGGAGACCTTCTGCAGCTCCCTGGT 657
Db 121 AlaTyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuVal 140
QY 658 ACACAGCAACATCCCAACGCTTTTCTCCATGCAGATGTGTGGAGCCGCTTGCCCGTT 717
Db 141 ThrGlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProVal 160
QY 718 GCTGGATCTGGACCAACGAGGTAGTCTGTCTGGGTGGAAATGAACCAAGTTTGAT 777
Db 161 AlaGlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyr 180
QY 778 AAAGGAGACATCTGGTATACCCCTTATTAAGGAAGAGTGTACTACACATACAAATCTGT 837
Db 181 LysGlyAspIleThrProLysGlyGluTrpTyrTyrGlnIleGluIleLeu 200
QY 838 AAATTGGAATTTGGAGCCCAAGCCCTTAATCTGAGCTCAGAGAGTATTAACGACAGAG 897
Db 201 LysLeuGluIleGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLys 220
QY 898 GCATCTGTGACAGTGGCACCAGCTGTGCGCTGCCGCCAGAGGTGTTTGCATCGGTG 957
Db 221 AlaIleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaVal 240
QY 958 GTGGAAGCTGTGGCCCGCATCTCTGATTCAGAAATCTCTGATGGTTCTTSGACTGG 1017
Db 241 ValGluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGly 260
QY 1018 TCCCAGCTGGCTGGGAGCAATTCGGAACACACCTTGGTCTTACTTCCCTAAATCTCC 1077
Db 261 SerGlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSer 280
QY 1078 ATCTACTGAGAGATGAGAACTCCAGAGCTCATTCCTGATCACAATCTGCTCAGATT 1137
Db 281 IleTyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeu 300
QY 1138 TACATTCAGCCCATGATGGGGCGCCCTGAATTAAGATTATGCAATTCGGCAATTTC 1197
Db 301 TyrIleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSer 320
QY 1198 CCATCCACAATCCGCTGTGATCGGTCCAGCGGTATGAGGGCTTCTACGTCATCTTC 1257
Db 321 ProSerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePhe 340
QY 1258 GACAGAGCCGAGAGGGTGGCTTCCGAGCGAGCCCTGTGCAGAAATTCAGGTGCT 1317
Db 341 AspArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAla 360
QY 1318 GCAGTGTCTGAATTTCCGGGCTTTCTCAACAGAGAGATGACCCAGCAACTGTGTCGCC 1377
Db 361 AlaValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnLysValPro 380
QY 1378 GCTCAGCTCTTTCGAGCGAGCCCAATTTTGTGGATTGTCTCTATGCGCTCATGAGCGTCTGT 1437
```

```
Db 381 AlaGlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCys 400
QY 1438 GGAGCCATCTCTCTTCTTAATCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1497
Db 401 GlyAlaIleLeuLeuValIleValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 420
QY 1498 CCCCGTGCACCTGAGTGGTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1554
Db 421 ProArgProGluValValAsnAspLysSerLeuValArgHisArgTrpLys 439

RESULT 6
Q8N2D4 PRELIMINARY; PRT; 423 AA.
ID Q8N2D4 Q8N2D4; 01-OCT-2002 (TREMBlrel. 22, Created)
AC Q8N2D4; 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein OVARC1000363.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Ovarian carcinoma;
RA Ota T., Nishikawa T., Suzuki Y., Kawai-Hio Y., Hayashi K., Ishii S.,
RA Saito K., Yamamoto J., Wakamatsu A., Nagai T., Nakamura Y.,
RA Nagahari K., Sugano S., Isogai T.;
RT "HKI human cDNA sequencing project.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK075539; BAC11682.1;
DR InterPro: IPR001461; AspproteaseA1.
DR Pfam: PF00026; asp; 2.
DR PRINTS: PR00792; PEPFIN.
DR PROSITE: PS00141; ASP_PROTEASE; 2.
KW Hypothetical protein.
SQ SEQUENCE 423 AA; 46457 MW; 4D4839F2ED9C2CE1 CRC64;

Alignment Scores:
Pred. No.: 2,52e-160 Length: 423
Score: 2184.00 Matches: 420
Percent Similarity: 99.53% Conservative: 1
Best Local Similarity: 99.29% Mismatches: 2
Query Match: 65.41% Indels: 0
Gaps: 4

US-09-806-194-1 (1-1804) x Q8N2D4 (1-423)
```

```
QY 286 ATGCTGATCGGACCCCGCCAGAGCTACAGATTCTCGTTGACACTGGAAGCAGTAAC 345
Db 1 MetLeuIleGlyThrProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsn 20
QY 346 TTTGCCGTGGCAGGACCCCGCCTCTCTACATAGACAGCTACTTTCACACAGAGAGTCT 405
Db 2i PheAlaValAlaGlyThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSer 40
QY 406 AGCACAATACCCCTCCAGGGCTTTGACGTCACAGTGAAGTACACACAGGAAGCTGGACG 465
Db 41 SerThrTyrArgSerLysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThr 60
QY 466 GGCTTCGTTGGGGAAGACCTCGTCACCATCCCAAGGCTTCAATACTCTTCTTTCTTGT 525
Db 61 GlyPheValGlyGluAspLeuValThrIleProLysGlyPheAsnThrSerPheLeuVal 80
QY 526 AACATTGCCACTATTTTGAATCAGAGAAATTTCTTTTGGCCTGGGATTAATGAATGGA 585
Db 81 AsnIleAlaThrIlePheGluSerGlyAsnPhePheLeuProGlyIleGlnTrpAsnGly 100
QY 586 ATACTTGGCTAGCTTATGCCACACTTGCACAGCCATCAAGTCTCTCTGAGACCTTCTTC 645
Db 101 IleLeuGlyLeuAlaTyrAlaThrLeuAlaLysProSerSerLeuGluThrPhePhe 120
```

[illegible]

```
QY 601 TATGCCACACTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGAGCTCCCTGGTACA 660
Db 201 TyrAlaThrLeuAlaIysProSerSerLeuGluThrPhePheAspSerLeuValThr 220
QY 661 CAAGCAACATCCCAAGCTTTTCTCCATGCAGATGTGTGGACCGGCTTGGCCCTTCT 720
Db 221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240
QY 721 GCATCTGGACCAACGGAGGTAGTCTCTCTGGTGGGAATTAACCAACTTGTATAAA 780
Db 241 GlySerGlyThrAsnGlySerLeuValLeuGlyGlyLeuProSerLeuTyrIys 260
QY 781 GGAGACATCTGTATACCCCTATTAGGAAGAGTGTCTACACAGATAGAAATTCGAAA 840
Db 261 GlyAspIleTrpIleThrProIleIysGluIleTyrTyrGluIleLeuIys 280
QY 841 TTGGAAATTTGGAGGCCAAGCCCTTAATCTGGACTGCAGAGATATAACGCAGCAAGGCC 900
Db 281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspIysAla 300
QY 901 ATCTGGACAGTGGCACCACCGTGTGGCCCTGCCCCAGAGGTGTTTCATCGCGTGTG 960
Db 301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValAl 320
QY 961 GAAGCTGGCCCGCGCATCTCTGATTCAGAAATCTCTGATCGTTTCTGGACTGGGTCC 1020
Db 321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340
QY 1021 CAGCTGGCTGTGACGAAATTCGGAACACCTTGGTCTTACTTCCCTAAATCTCCATC 1080
Db 341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProIysIleSerIle 360
QY 1081 TACCTGAGAGATGAGAACTCCAGCAGGTGATTCCTGATCACATCTGCTCAGCTTAC 1140
Db 361 TyrLeuArgAspGlnAsnSerSerArgSerPheArgIleThrIleLeuProGln----- 378
QY 1141 ATTCAGCCCATGATGGGGCGCGCTGAATTTATGAATGTATCCGATTTCGGCATTTCC 1200
Db 378 ----- 378
QY 1201 TCACAAATGCCTGGTGATCGGTGCCAGGTGATGGAGGGCTTACGTCATCTCGAC 1260
Db 378 ----- 378
QY 1261 AGAGCCCAAGAGGTGGGCTTCGACGAGCCCTGTGACAGAAATTCAGGTCTCGCA 1320
Db 379 -----Lys-LeuArgValLeuG 384
QY 1321 GTGCTGAAATTTCCGGGCTTTCTCAACAGAGGATG 1357
Db 384 nCysLeuLysPheProGlyLeuSerGlnArgMet 396
```

RESULT 8

```
Q9R1P7
ID Q9R1P7 PRELIMINARY: PRT: 255 AA.
AC Q9R1P7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2003 (TrEMBLrel. 23, Last annotation update)
DE Aspartyl protease (Fragment).
GN BACE2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Accarino M., Fumagalli P., Taramelli R., Ottolenghi S.;
RT "Cloning of a gene from chromosome 21 Down Region encoding a potential
RT transmembrane protease."
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF051150; AAD45964.1; -.
DR MEROPS: A01.041; -.

```

```
MGD; MGI:1860440; Bace2.
InterPro; IPR001969; Aspartic_acid_protease_site.
PROSITE; PS00141; ASP_PROTEASE; 1.
KW Protease.
FT NON_TER 1
SQ SEQUENCE 255 AA; 28685 MW; 53DE317815996D63 CRC64;

Alignment Scores:
Pred. No.: 7,61e-88 Length: 255
Score: 1246.00 Matches: 232
Percent Similarity: 95.69% Conservative: 12
Best Local Similarity: 90.98% Mismatches: 11
Query Match: 37.32% Indels: 0
DB: 11 Gaps: 0

US-09-806-194-1 (1-1804) x Q9R1P7 (1-255)
QY 790 TGTATACCCCTATTAGGAAGAGTGGTACTACACAGATAGAAATTCGAAATTTGAAATT 849
Db 1 TrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLysLeuGluIle 20
QY 850 GGAGGCCAAACCTTAACTCTGGACTGCAGAGATATAACGCAGACAGGCCATCGTGAC 909
Db 21 GlyGlyGlnAsnLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAlaIleValAsp 40
QY 910 AGTGCACACAGCTGCTGCCTGCCCCAGAGGTGTTGTATCGGTGGTGAAGCTGTG 969
Db 41 SerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValAlaGluAlaVal 60
QY 970 GCCCGCGCATCTCTGATTCACAGAAATCTCTGATGGTTTCTGGACTGGGTCCAGCTGGCG 1029
Db 61 AlaArgThrSerLeuIleProGluPheSerAspGlyPheTrpThrGlyAlaGlnLeuAla 80
QY 1030 TCTGGAGCAATTCGGAACACCTTGGTCTTACTTCCCTAAATCTCCATCTACCTGAGA 1089
Db 81 CysTrpThrAsnSerGluThrProTrpAlaTyrPheProLysIleSerIleTyrLeuArg 100
QY 1090 GATGAGAATCTCCAGCAGTTCCTCGTATCAATCTCCCTGCTCAGCTTTACATTCAGCCC 1149
Db 101 AspGluAsnAlaSerArgSerPheArgIleThrIleLeuProGlnLeuTyrIleGlnPro 120
QY 1150 ATGATGGGGCGCGCTGAAATTTATGAATTTACCGATTTCCGCATTTCCCATCCACAAAT 1209
Db 121 MetMetGlyAlaGlyPheAsnTyrGluCystyrArgPheGlyIleSerSerThrAsn 140
QY 1210 CGCTGGTGATCGGTGCCACCGTGTATGAGGCTTCTACGTCATCTTCGACAGAGCCGAG 1269
Db 141 AlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValValPheAspArgAlaGln 160
QY 1270 AAGAGGTGGGCTTCGACGAGCCCTGTGCAGAAATTTGCAGGTGCTGCAGTGTCTGAA 1329
Db 161 ArgArgValGlyPheAlaValSerProCysAlaGluIleGluGlyThrThrValSerGlu 180
QY 1330 ATTTCCGGGCGCTTCTCAACAGAGATGATGCCAGCAACTGTGTCCCGCTCAGTCTTTG 1389
Db 181 IleSerGlyProPheSerThrGluAspIleAlaSerAsnCysValProAlaGlnAlaLeu 200
QY 1390 ACCGAGCCCATTTGTGGATTGTCTCTATCGCTATGAGCGCTGTGGAGCCATCTC 1449
Db 201 AsnGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGlyAlaIleLeu 220
QY 1450 CTTGTCTTAATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1509
Db 221 LeuValLeuIleLeuLeuLeuLeuValProLeuHisCysArgHisAlaProArgAspPro 240
QY 1510 CAGGTGCTCAATGATGAGTCTCTCTCTGCTCAGACATCGCTGAAA 1554
Db 241 GluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 255

RESULT 9
Q9ULS1
IC Q9ULS1 PRELIMINARY: PRT: 532 AA.
AC Q9ULS1;

```

Db	164	ThrGlnGlyLysTrpGluGlyLeuGlyThrAspLeuValSerIleProHisGlyPro	183
QY	508	ATAATCTCTTTCTTCAACATGCCACTATTTTGAATCAGAGAAATTTCTTTTGCT	567
Db	184	AsnValThrValArgAlaAsnIleAlaAlaIleThrGluSerAspLysPheIleAsn	203
QY	568	GGGATTAAATGGAAATGAAATCTTGGCTTAGCTTTATGCCACACTTGCACAGCCATCAAGT	627
Db	204	GlySerAsnTrpGluGlyIleLeuGlyLeuAlaTyraIleAlaArgProAsp	223
QY	628	TCTCTGGAGACTTCTTCGACTCCCTGGTGACACAAACATCCCAAGCTTTCTTCC	687
Db	224	SerLeuGluProPhePheAspSerLeuValLysGlnThrHisValProAsnLeuPheSer	243
QY	688	ATGCAGATGTCTGGAGCGGCTTGCOCGTTCGGTAICT-----GGGACCAACGGA	738
Db	244	LeuGlnLeuCysGlyAlaGlyPheProLeuAsnGlnSerGluValLeuAlaSerValGly	263
QY	739	GCTAGCTGTCTTCGGTGGAAATGAACCAAGTTGTATAAAGGAGACATCTGGTATACC	798
Db	264	GlySerMetIleIleGlyGlyIleAspHisSerLeuTyThrGlySerLeuTrpTrpThr	283
QY	799	CCTATTAAAGGAAGTGGTACTACACAGATAGAAATTTCTGAAATTTGGAGGCCAA	858
Db	284	ProIleArgArgGluTyrTyrGluValIleIleValArgValGluIleAsnGlyGln	303
QY	859	AGCCTTAATCTGGACTCGACAGAGTATAACGACAGCAAGCCATCGTGACAGTGGCACC	918
Db	304	AspLeuLysMetAspCysLysGluTyraTyrAspLysSerIleValAspSerGlyThr	323
QY	919	AGCCTGTCTGCCCTGCCCAAGAGTGTATTGATCGGTGGTGGAAAGCTGTGCCCGCGCA	978
Db	324	ThrAsnLeuArgLeuProLysLysValPheGluAlaValLysSerIleLysAlaAla	343
QY	979	ICTCTGATTCAGAAATCTCTGATGGTTTCGACTGGGTCCAGCTGCCTGCCTGGGACG	1038
Db	344	SerSerThrGluLysPheProAspGlyPheTrpLeuGlyGluLeuValCysTrpGln	363
QY	1039	AATTCGGAAACACCTCTGTCTTACTTCCTTAAATCTCCATCTACTCAGAGATGAGAAC	1098
Db	364	AlaGlyThrThrProTrpAsnIlePheProValIleSerLeuTyLeuMetGlyGluVal	383
QY	1099	TCCACAGGTCATTCGGTATCATCAATCTGCCTCAGCTTTTACATTCAGCCCATGATGGG	1158
Db	384	ThrAsnGlnSerPheArgIleThrIleLeuProGlnGlnTyLeuArgProValGluAsp	403
QY	1159	GCGCGCTCGAATTAI---GAAATGTACCAATTCGGCATTTCCCATTCACAAATSCGCTG	1215
Db	404	ValAlaThrSerGlnAspAspCysTyrrLysPheAlaIleSerGlnSerThrGlyThr	423
QY	1216	GTGATCGTGGCCAGGTGATGGAGGGCTTCTACGTCACTTCGACAGCCCGAGAAGAGG	1275
Db	424	ValMetGlyAlaValIleMetGluGlyPheTyrrValValPheAspArgAlaArgLysArg	443
QY	1276	GTGGCTTCGACGAGCGCCCTCTGTCAGAAATTCGAGGTGCTGCATGTCCTCAAAATTC	1335
Db	444	IleGlyPheAlaValSerAlaCysHisValHisAspGluPheArgThrAlaAlaValGlu	463
QY	1336	GGGCGCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCGCTCAGTCTTTGACGAG	1395
Db	464	GlyProPheValThrLeuAspMetGluAspCysGlyTyrrAsnIleProGlnThrAspGlu	483
QY	1396	CCCATTTTGTGATGTGCTTCTATGCCCTCATGAGCGTCTGTGGAGCCATCTCTTTGTC	1455
Db	484	SerThrLeuMetThrIleAlaTyrrValMetAlaIleCys---AlaLeuPheMetLeu	502
QY	1456	TTAAATGTCCTCTGCTGCTGCCCTCCGGTGTACAGCGT-----CGCCCCGCTCACCT	1509
Db	503	ProLeuCysLeuMetValCysGlnTrpArgCysLeuArgCysLeuArgGlnGlnHisasp	522
QY	1510	GAGGTGCTCAATGATGATGCTCTCTG	1536
Db	523	AspPheAlaAspAspIleSerLeuLeu	531

RA Strausberg R.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBSJ databases.
 DR EMBL: BC036084; AAH36084.1; -; 768595CF5517EFB7 CRC64;
 SQ SEQUENCE 501 AA: 55824 MW: 768595CF5517EFB7 CRC64;

Alignment Scores:

Pred. No.: 4,59e-82 Length: 501
 Score: 1172.50 Matches: 239
 Percent Similarity: 61.85% Conservative: 82
 Best Local Similarity: 46.05% Mismatches: 165
 Query Match: 35.12% Indels: 33
 DB: 4 Gaps: 9

US-09-806-194-1 (1-1804) x Q8IYC8 (1-501)

QY 19 CGCGTGTGCTGCTCTGCTGSCCCAGTGGCTCTGCGCGCGCGGAGCTGGCCGC 78
 DB 2 AlaGlnAlaLeuProTrpLeuLeuLeuTrpMet-----GlyAlaGlyValLeuPro 16
 QY 79 CGGCCCTTCAG-----CTGCCCTCGCGGTGGCGCGGCCACGACCGC 123
 DB 19 AlaHisGlyThrGlnHisGlyLeuArgLeuProLeuArgSerGlyLeuGly----- 35
 QY 124 GTAGTGGCCGCCCGGGACCGGGACCCCTGCGGAGCCCGACCGCGGTGGCG 183
 DB 36 -----GlyAlaProLeu-----GlyLeuArg 42
 QY 184 CTGCGCCCTGGAG-----CTGCGCTGCGCGCGCGCGCGCAACTTCTTGGC 237
 DB 43 LeuProArgGluThrAspGluGluProGluGluProGlyArgArgGlySerPheValGlu 62
 QY 238 ATGTAGAACACCTGCAGGGGACTTGGCGCGCGCTACTACCTGGAGATCTGATCGG 297
 DB 63 MetValAspAsnLeuArgGlyLysSerGlyGlnGlyTyrIrrValGluMetThrValGly 82
 QY 258 ACCCCCGCGAGACCTACACATCTCTGTCACACTGGAGACGACGTAACCTTGGCG 357
 DB 83 SerProProGlnThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaValGly 102
 QY 358 GGAACCCCGCACTCTACATAGACAGCACTTTCACACAGAGAGTCTAGCACATACCG 417
 DB 103 AlaAlaProHisProPheLeuHisArgTyrTyrGlnArgGlnLeuPheSerIrrTyrArg 122
 QY 418 TCCAGGCGCTTGACGTACAGTGAAGTACACAGAGAGCTGGAGCGGCTCTCTGG 477
 DB 123 AspLeuArgLysGlyValTyrValProTyrThrGlnGlyLysTrpGluGlyGluLeuGly 142
 QY 478 GAAGACCTCGTACACATCCCAAGGCTTCAATCTCTTTCGTCAACATTCGCACT 517
 DB 143 ThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAlaAla 162
 QY 538 ATTTTGAATCAGAAATTTCTTTTGGCTGGGATTAATGGAATGGAATCTGGGCTA 597
 DB 163 IleThrGluSerAspLysPhePheIleAsnGlySerAsnTrpGluGlyIleLeuGlyLeu 182
 QY 598 GCTTATGACACATGCCAGGCATCACTCTCTGGAGACCTCTCTGACCTCCCTGGT 657
 DB 183 AlaTyrAlaGluIleAlaArgProAspSerLeuGluProPhePheAspSerLeuVal 202
 QY 658 ACACAGCAACATCCCAAGCTTTCTCCATGCAATGTGTGGAGCGGCTGCGCCCTT 717
 DB 203 LysGlnThrHisValProAsnLeuPheSerLeuGlnLeuCysGlyAlaGlyPheProLeu 222
 QY 718 GCTGGATCT-----GGACCAACGAGGAGTACTCTTGTGGTGGAAATGAACA 768
 DB 223 AsnGlnSerGluValLeuAlaSerValGlyGlySerMetIleIleGlyGlyLeuAspHis 242
 QY 769 AGTTTGTATAGGACACATCTGGTATACCCCTATTAGGAGAGCTGTACTACAGATA 828
 DB 243 SerLeuTyrThrGlySerLeuTrpTyrThrProIleArgArgGluTrpTyrGluVal 262
 QY 829 GAAATCTGAAATCGAAATCGAGGCCAAAGCCCTTAATCTGGACTCGAGAGATATAAC 888
 |||||:|||||:||||| ||||| ||| :|||||:|||||

DB 263 IleIleValArgValGluIleAsnGlyGlnAspLeuLysMetAspCysLysGluTyrAsn 282
 QY 889 GCACACAAGCCCATCTGGACAGTGCACACACAGCTCTGCGCTGCCCAAGAAGGTGTTT 948
 DB 283 TyrAspLysSerIleValaspSerGlyThrThrAsnLeuArgLeuProLysLysValPhe 302
 QY 949 GATGCGGTGGGAAGCTGTGCCCGCATCTCTGATTCCAGAAATCTCTGATGTTTC 1008
 DB 303 GluAlaAlaValLysSerIleLysAlaAlaSerSerThrGluLysPheProaspGlyPhe 322
 QY 1009 TGGACTGGTCCCGAGCTGGCGTCTGGACCAATTCGGAACACCTTGGCTTACTTCCCT 1068
 DB 323 TrpLeuGlyGluGlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnIlePhePro 342
 QY 1069 AAAATCTCCATCTACCTGACAGATGAGAACTCCAGCAGGTCATTCGGTATCACAATCTCG 1128
 DB 343 ValIleSerLeuTyrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIleLeu 362
 QY 1129 CTTACGCTTTTACATTCAGCCCATGATGGGGCGCGCTGAATTA---GAATGTTACCGA 1185
 DB 363 ProGlnGlnTyrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTyrLys 382
 QY 1186 TTCGGCATTTCCCATCCACAAATGCGCTGATCGTGGTCCACGGTGATGGAGGCTTC 1245
 DB 383 PheAlaIleSerGlnSerThrGlyThrValMetGlyAlaValIleMetGluGlyPhe 402
 QY 1246 TAGCTCATCTTCACACAGAGCGGTGGCTTCGACGAGCGCCCTGTGCAGAA 1305
 DB 403 TyrValValPheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHisVal 422
 QY 1306 ATTCAGAGTCTCGAGTGTCTGAAATTCGGGCCCTTTCTCAACAGAGGATGAGCCAGC 1365
 DB 423 HisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetG:uasp 442
 QY 1366 AACTGTCTCCCGCTCAGTCTTTCAGGAGCCCATTTTGGGATGTGTGCTATGGCTC 1425
 DB 443 CysGlyTyrAsnIleProGlnThrAspGluSerThrLeuMetThrIleAlaTyrValMet 462
 QY 1426 ATGAGCTGTGTGGAGCCATCTCTTGTCTTAATGCTGCTGCTGCTGCTGCTGCTGCGG 1485
 DB 463 AlaAlaIleCys---AlaLeuPheMetLeuProLeuCysLeuMetValCysGlnTyrArg 481
 QY 1486 TGCAGCGT-----CGCCCGCTGACCCCTGAGGTGCTCAATGATGATGCTCTCTG 1536
 DB 482 CysLeuArgCysLeuArgGlnGlnHisAspPheAlaAspAspIleSerLeuLeu 500

RESULT 13

Q8C4F4 PRELIMINARY; PRT; 467 AA.

AC Q8C4F4: 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Beta-site APP cleaving enzyme.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
 RA MEDLINE=22354683; PubMed=1246851;
 RA The FANTOM Consortium,
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RI 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 DR EMBL; AK082317; BAC38462.1; -;
 SQ SEQUENCE 467 AA; 52063 MW; 31AB674FF1843652 CRC64;

Alignment Scores:

Pred. No.: 1.6e-72 Length: 467
 Score: 1049.00 Matches: 215


```

Percent Similarity: 56.70%      Conservative: 77
Best Local Similarity: 41.75%    Mismatches: 163
Query Match: 31.42%             Indels: 60
DB: 11                           Gaps: 8

us-09-806-194-1 (1-1804) x 08C4F4 (1-467)

QY 25 CTGCTGCTCTGTCGCCAGTGGCTCTGCGCGCGCGCGCGAGCTGGCCCGCCGCGCC 84
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 1 MetAlaProAlaLeuHisTrpLeuLeuLeuLeuTrpValGlySerGlyMetLeuProAlaGln 20
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111
QY 85 TTCACG-----CTGCCCTCCGGGTGGCGCGCGCCAGCAACCGGTAGT 129
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 21 GlyThrHisLeuGlyLeuArgLeuProLeuArgSerGlyLeuAla----- 35
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111
QY 130 GCGCCACCCCGCGGAGCCCGCTCGCGAGCGCGCGCGCGCTGGCGCTCGCC 189
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 36 GlyProProLeuGlyLeuArgLeuProArgGluThrAspGluGluSer----- 51
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111
QY 190 CTGGAGCTCCCTGGGTCCCGCGCGCGCGCGCGCAACTTCTTGCCCATGTCAGAAC 249
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 52 -----GluGLuProGlyArgArgGlySerPheValGluMetValAspAsn 66
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111
QY 250 CTGACGGGGACTCTGGCGCGCGCTACTACCTGGAGATGCTGATCGGACCGCCGCGAG 309
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 67 LeuArgGlySerGlyGlnGlyTrpValGluMetThrValGlySerProProGln 86
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111
QY 310 AAGCTACAGATTCTGTTGACACTGAAGAGTAACTTTCGGCTGCGCAGAACCCGCGAC 369
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 87 ThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaValGlyAlaAlaProHis 106
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111
QY 370 TCCTACATACAGCTACTTTCACACAGAGAGTCTAGCACATACCGCTCCAGAGGCTT 429
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 107 ProPheLeuHisArgTrpGlnArgGlnLeuSerSerThrTrpArgAspLeuArgLys 126
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111
QY 430 GAGCTCACAGTGAAGTACACACAAAGGAGTCTGGCGGCTTCCTGGGAAGACCTGTC 489
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 127 GlyValTrpValProTrpThrGlnGlyLysTrpGluGlyGluLeuLysThrAspLeuVal 146
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111
QY 490 ACCATCCCAAGCGCTTCATCTCTTCTTCCTCAACATTCGCACTTTCGATCA 549
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 147 SerIleProHisGlyProAsnValThrValArgAlaAsnIleAlaAlaIleCThrGluSer 166
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111
QY 550 GAGAAATTTCTTTGGCTGGGATTAATGGAATGGAATCTTGGCTAGCTATGTCACA 609
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 167 AspLysPheIleAsnGlySerAsnTrpGluGlyIleLeuGlyLeuAlaTrpArgLys 186
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111
QY 610 CTGGCCAAAGCCATCAAGTTCTCTGGAGACTTCTTCGACTCCCTGGTGACACAGCAAC 669
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 187 IleAlaArgProAspAspSerLeuGluProPhePheAspSerLeuValLysGlnThrHis 206
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111
QY 670 ATCCCAAGCTTTCTCCATGCGAGTGTGGAGCGGCTTCCCGCT-----GCT 720
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 207 IleProAsnIlePheSerLeuGlnLeuCysGlyAlaGlyPheProLeuAsnGlnThrGlu 226
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111
QY 721 GGATCTGGGACCAACGAGTACTTCTTCTGGTGGATTAATGGAATGGAATTTGATAAA 780
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 227 AlaLeuAlaSerValGlySerMetIleIleGlyIleAspHisSerLeuTrpThr 246
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111
QY 781 GGAGACATCTGGTATACCCCTATTAGGAAGAGTGTGTACTACAGATGAATTTCTGAAA 840
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 247 GlySerLeuTrpTrpThrProIleArgArgGluTrpTrpTrpGluValIleValArg 266
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111
QY 841 TTGGAATTTGGAGCCAAAGCTTAATCTGACTGTCAGAGATTAACCCACAGAGGCC 900
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 267 ValGluIleAsnGlnAspLeuLysMetAspCysLysGlu----- 280
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111
QY 901 ATCGTGGACAGTGGCACCACGCTGCTGCGCGCTGCCCGGAGGTGTTTGTGCGGTG 960
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 280 ----- 280
QY 961 GAAGCTGTGGCGCGCGCATCTCTGATTCAGAAATCTCTGATGGTTTCGGACTGGGTCC 1020
   ::::: 111 111 111 111 111 111 111 111 111 111 111 111 111

```

RT *Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.*;
 RL Nature 420:563-573(2002).
 RN [3]

RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Brain;
 RX MEDLINE=21085660; PubMed=11217851;
 RA RIKEN FANTOM Consortium;
 RT *Functional annotation of a full-length mouse cDNA collection.*;
 RL Nature 403:685-690(2001).
 RN [4]

RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Brain;
 RX MEDLINE=99279253; PubMed=10349636;
 RA Carninci P., Hayashizaki Y.;
 RT *High-efficiency full-length cDNA cloning.*;
 RL Meth. Enzymol. 303:19-44(1999).
 RN [5]

RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Brain;
 RX MEDLINE=20499374; PubMed=11042159;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT *Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes.*;
 RL Genome Res. 10:1617-1630(2000).
 RN [6]

RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Brain;
 RX MEDLINE=20530913; PubMed=11078861;
 RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwara S., Inoue K., Iogawa Y., Iwawa M., Ohara E., Wataniki M.,
 RA Itoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsunura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT *RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer.*;
 RL Genome Res. 10:1757-1771(2000).
 DR EMBL: AK014390; BAB29317.2; .
 FT NON_TER 1 1
 SQ SEQUENCE 267 AA; 30333 MW; 9413EB4530AB63B0 CRC64;

Alignment Scores:
 Pred. No.: 5,68e-42 Length: 267
 Score: 653.00 Matches: 121
 Percent Similarity: 66.29% Conservative: 56
 Best Local Similarity: 45.32% Mismatches: 86
 Query Match: 19.56% Indels: 4
 DB: 11 Gaps: 3

US-09-806-194-1 (1-1804) x Q9CU05 (1-267)

QY 745 CTGTCTTGGTGGATGACCAAGTTGTATAAGAGGACATCTGGIATACCCCTATT 804
 Db 1 MetilleleGlyGlyleAspHisSerLeuThrGlySerLeuThrThrProile 20

QY 805 AAGGAAGTGGTACTACCATAGAAATTCIGAAATGGAAATGGAGCCAAAGCCTI 864
 Db 21 ArgArgGluThrTyThrGluValleleValArgValGluIleAsnGlyGluAspLeu 40

QY 865 AATCTGGACTCAGAGAGTATACCCAGACAGGACATCTGTGGACAGTGGCAGCCAGCTG 924
 Db 41 LysMetAspCysLysGluTyThrAsnTyThrAspLysSerleleValAspSerGlyThrThrAsu 60

QY 925 CTGGCCCTGCCAGACAGTGTGTATCGCGTGGTGGAGCTGTGGCCCGGCGATCTCTG 984
 Db 61 LeuArgLeuProLysLysValPheGluAlaValLysSerleleLysAlaAlaSerSer 80

QY 985 ATCCAGAAATCTCTGATGGTTCTGGACCTGGCCAGCTGGCGTGGTGGACCAATTCG 1044
 Db 81 ThrGluLysPheProAspGlyThrleuPheGlyGluGlnLeuValCysTrpGlnAlaGly 100

QY 1045 GAACACACCTTGGTCTTACTTCCCTAAATCTCCATCTACCTGAGAGATCAGAACTCCACG 1104
 Db 101 ThrThrProTrpAsnIlePheProValIleSerLeuTyLeuMetGlyGluValThrAsn 120

QY 1105 AGTCATTCCTGATCACAATCTGCTCAGCTTTCACATTCAGCCATCAGCCATGGGGCCGCG 1164
 Db 121 GlnSerPheArgIleThrleLeuProGlnGlnTyLeuArgProValGluAspValAla 140

QY 1165 CTGAATAT---GAATGTTACCGATTCGGCATTTCCCATCCCAAAATCCGCTGGTGATC 1221
 Db 141 ThrSerGlnAspCysTyThrLysPheAlaValSerGlnSerSerThrGlyThrValMet 160

QY 1222 GTGCCACGCTGATGAGCGCTTCTACGTCATCTTCGACAGCCAGAGAGGGTGGGC 1281
 Db 161 GlyAlaValIleMetGluGlyPheTyValValPheAspArgAlaArgLysArgIleGly 180

QY 1282 TTGGCAGCGAGCCCTGTGCAGAAATTCGAGTGTCTGCTGAAATTTCCGGGCT 1341
 Db 181 PheAlaValSerAlaCysHisValHisAspGluPheArgThrAlaAlaValGluGlyPro 200

QY 1342 TTCTCAACAGAGATGTAGCCCAACTGTGTCCCGCTCAGCTCTTGAGCGAGCCCAT 1401
 Db 201 PheValThrAlaAspMetGluAspCysGlyTyThrAsnIleProGlnThrAspGluSerThr 220

QY 1402 TTGTGATTTGTCTTATCGCTCATGAGCGCTGTGTGGAGCCATCTCTCTCTTAATC 1461
 Db 221 LeuMetThrIleAlaTyValMetAlaAlaIleCys---AlaLeuPheMetLeuProLeu 239

QY 1462 GTCTCTGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1515
 Db 240 CysLeuMetValCysGlnTrpArgCysLeuArgCysLeuArgHisGlnHisAspAspPhe 259

QY 1516 GTCAATGATGATGCTCTCTG 1536
 Db 260 AlaAspAspLeuSerLeuLeu 266

RESULT 15
 Q9PD02 PRELIMINARY; PRT; 213 AA.

DI 01-OCT-2000 (TREMBlrel. 15, Created)
 DI 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DI 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE HSPC104 (Fragment).

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

RP SEQUENCE FROM N.A.
 RC TISSUE=Umbilical cord blood;
 RA Zhang O.H., Ye M., Zhou J., Shen Y., Wu X.Y., Guan Z.Q., Wang L.,
 RA Fan H.Y., Mao Y.F., Dai M., Huang Q.H., Chen S.J., Chen Z.;
 RT "Human partial CDS cloned from cd34+ stem cells.";
 RL Submitted (MAY-1999) to the FMBL/GenBank/DBJ databases.
 DR EMBL: AF161367; AAF28927.1; .
 DR InterPro: IPR001461; AspproteaseA1.

DR Pfam: PF00026; asp; 1.
 DR NON_TER 1 1
 SQ SEQUENCE 213 AA; 24338 MW; EC9D3FA31CFA835C CRC64;

Alignment Scores:
 Pred. No.: 3,75e-27 Length: 213
 Score: 461.00 Matches: 84
 Percent Similarity: 68.10% Conservative: 27
 Best Local Similarity: 51.53% Mismatches: 42
 Query Match: 13.81% Indels: 10
 DB: 4 Gaps: 1

US-09-806-194-1 (1-1804) x Q9PD02 (1-213)

QY 754 GTGGGAATGAACCAAGTTGTATAAGAGGAGACATCTGGTATACCCCTATTAGGAAGAG 813

```
Db      |||||
1 GlycylIleAspHisSerLeuTyrThrGlySerLeuTyrThrProIleArgGlu 20
QY      814 TGGTACTACGAGATAGAAATTCGAAATTCGAAATTCGAGGCCAAAGCCITTAATCTGGAC 873
Db      |||||
21 TrpTyrTyrGluValIleIleValA:GValGluIleAsnGlyGlnAspLeuLysMetAsp 40
QY      874 TGCAGAGAGTATAACGACAGAGGCCATCGTGCACAGTGGCACACACGCTGCTGGCGCTG 933
Db      |||||
41 CysLysGluTyrAsnTyrAspLysSerIleValAspSerGlyThrThrAsnLeuArgLeu 60
QY      934 CCCAGAGGTGTTGATCGGTGGTGGAGCTGTGGCCCGGCACTCTCTGATTCAGAA 993
Db      |||||
61 ProLysLysValPheGluAlaAlaValLysSerIleLysAlaAlaSerSerThrGluLys 80
QY      994 TTCTCTGATGTTCTGCGACTGGTCCGACGTGGCTGGTGGACGAATTCGGAAACACCT 1053
Db      |||||
81 PheProAspGlyPheIrpLeuGlyGluGlnLeuValCysTrpGlnAlaGlyThrThrPfc 100
QY      1054 TGGTCTTACTTCCTTAAATCTCCATCTACCTGAGAGATGAGAACTCCAGCAGGTCATTC 1113
Db      |||||
101 TrpAsnIlePheProValIleSerLeuTyrLeuMetGlyGluValThrAsnGlnSerPhe 120
QY      1114 CGTATCACAACTCCGCTCAGCTTTACATTCACCCCATGATGGGGCCGCCCTGAATTAT 1173
Db      |||||
121 ArgIleThrIleLeuProGlnGlnTyrLeuArgPro----- 132
QY      1174 GAATGTTACCGATTCCGSCATITCCCATCCACAAATCCGCTGGTGTGATCGTCACAGGTS 1233
Db      |||||
133 -----TrpLysMetTrpPrcArgProLysThrThrValThrValCysHisLeuThrVal 150
QY      1234 ATGGAGGCC 1242
Db      |||||
151 IleHisGly 153
```

Search completed: September 25, 2003, 11:07:16
Job time : 204 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	2687	80.5	5.18	1	BAE2_HUMAN	Q9Y520	homo sapien
2	1187	35.5	501	1	BACE_RAT	P56819	rattus norv
3	1185	35.5	501	1	BACE_MOUSE	P56818	mus muscuiu
4	1178.5	35.3	501	1	BACE_HUMAN	P56817	homo sapien
5	363.5	10.9	377	1	PEPC_MACTU	P03955	macaca fusc
6	353	10.6	388	1	PEPC_HUMAN	P20142	homo sapien
7	351.5	10.5	388	1	PEPC_CALUA	P092D3	callithrix
8	324.5	9.7	394	1	PEPC_CAVPO	Q64411	cavia porce
9	320	9.6	402	1	RENI_MOUSE	P06281	mus muscuiu
10	315.5	9.4	396	1	CATD_CLUHA	Q9dex3	clupea hare
11	313.5	9.4	509	1	APR1_ORYSA	O42456	oryza sativ
12	313	9.4	392	1	PEPC_RAT	P04073	rattus norv
13	310	9.3	383	1	PEPC_CHICK	P16476	gallus gall
14	308.5	9.2	412	1	CATD_HUMAN	P18242	mus muscuiu
15	306.5	9.2	410	1	CATD_MOUSE	P00796	mus muscuiu
16	305.5	9.1	401	1	RENS_MOUSE	P24268	rattus norv
17	305	9.1	407	1	CATD_RAT	P56272	gadus morhu
18	302	9.0	424	1	PEP1_GADMO		


```

QY 481 GACCTGTCACATCCCAAGAGTTCATCTCTTCTGTCACATTCGCCACATTT 540
DB 161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180
QY 541 TTTGAATCAGAAATTTCTTTTGGCTGGGATTAATGGAATGAAACTTGGGCTACT 600
DB 181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200
QY 601 TATGCCACACTTGGCAAGCCATCAAGTCTCTGTGAGACCTTCTTGCATCTCCCTGGTACA 660
DB 201 TyrAlaThrLeuAlaLysProSerSerLeuGluThrPhePheAspSerLeuValThr 220
QY 661 CAAGCAATATCCCAAGAGTTCATCTCTGATGAGATGTGTGAGCGGCTGCCGCTTCTCT 720
DB 221 GluAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240
QY 721 GGATCTGGGACCAAGGAGTGTCTCTGGTGGATTAATGGACCAAGTTCGTATAAA 780
DB 241 GlySerGlyThrAsnGlyLysLeuValLeuGlyIleGluProSerLeuThrLys 260
QY 781 GGAGACATCTGTATACCCCTATTAAAGGAAGAGTGTACTACCAAGATAGAAATTCGTGAAA 840
DB 261 GlyAspIleTrpThrProIleLysGluGluTrpTyrGlnIleGluIleLeuLys 280
QY 841 TTGGAATTTGGAGGCCAAGAGCTTAATCTGGACTCCAGAGATATACGCAGACAGGCC 900
DB 281 LeuGluIleGlyGlnSerLeuAsnLeuAspCysArgGluTrpAsnAlaAspLysAla 300
QY 901 ATCGTGGACAGTGGACACACAGCTGTCTGCGCTGCCGCCAGAGGTTTTCATCGCGTGGT 960
DB 301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320
QY 961 GAAGTGTGGCCCGCGCATCTCTGATTCAGAAATTCCTGATGGTTCCTGGACTGGGTCC 1020
DB 321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340
QY 1021 CAGCTGGGTGTGGCAAGATTCGGAACACCTTGCTTACTTCCCTAAATCTCCATC 1080
DB 341 GlnLeuAlaCysTrpThrAsnSerGluThrProIleProIlePheProLysIleSerIle 360
QY 1081 TACCTGAGAGATGAGAATCTCCAGAGTTCATTCCTATCACAAATCTCCCTCAGGTTTAC 1140
DB 361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleIleThrIleLeuProGlnLeuLys 380
QY 1141 ATTCAGCCCATGATGGGCGCGCTGATTAATGATTAATGATTAATGATTAATGATTAATG 1200
DB 381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400
QY 1201 TCACAAAAGCGCTGTGTGATCGTCCAGCGGTGATGGAGGCTTCTACTCATCTCTGAC 1260
DB 401 SerThrAsnAlaLeuValIleLysAlaThrValMetGluGlyPheTyrValIlePheAsp 420
QY 1261 AGAGCCCAAGAGGTTGGGCTTCGACGAGCCGCTCTGACGAAATTCGAGGTCCTGCA 1320
DB 421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440
QY 1321 GTGCTGTAATTTCCGGGCTTCTTCACAGAGATGTAGCCAGCACTGTCTCCGCTCT 1380
DB 441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460
QY 1381 CAGTCTTTTCCAGGAGCCATTTTGTGGATGTGTCTATGCGCTCATGAGGCTCTGTGGA 1440
DB 461 GlnSerLeuSerGluProIleLeuThrIleValSerTyrAlaLeuMetSerValCysGly 480
QY 1441 GCCATCTCTCTTCTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
DB 481 AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 500
QY 1501 CGTGACCTGAGTGTCTCAATGATGAGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
DB 501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518

```

```

RESULT 2
BACE_RAT
ID BACE_RAT STANDARD; PRT: 501 AA.
AC P56819;
DI 30-MAY-2000 (Rel. 39, Created);
DI 30-MAY-2000 (Rel. 39, Last sequence update);
DI 28-FEB-2003 (Rel. 41, Last annotation update);
DE Beta-secretase precursor (EC 3.4.23.-) (Beta-site APP cleaving enzyme)
DE (beta-site amyloid precursor protein cleaving enzyme) (Aspartyl
DE protease 2) (Asp 2) (ASP2) (Membrane-associated aspartic protease 2)
DE (Memapsin-2).
GN BACE.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20002972; Pubmed=10531052;
RA Vassar R., Bennett B.D., Babu-Khan S., Kahn S., Mendiaz E.A.,
RA Denis P., Teplow D.B., Ross S., Amarante P., Loeloff R., Luo Y.,
RA Fisher S., Fuller J., Edenson S., Lile J., Jarosinski M.A.,
RA Biere A.L., Curran E., Burgess T., Louis J.-C., Collins F.,
RA Treanor J., Rogers G., Citron M.;
RA "Beta-secretase cleavage of Alzheimer's amyloid precursor protein by
RI the transmembrane aspartic protease BACE.";
RL Science 286:735-741(1999).
CC -!- FUNCTION: RESPONSIBLE FOR THE PROTEOLYTIC PROCESSING OF THE
CC AMYLOID PRECURSOR PROTEIN (APP). CLEAVES AT THE AMINO TERMINUS OF
CC THE A-BETA PEPTIDE SEQUENCE, BETWEEN RESIDUES 671 AND 672 OF APP.
CC LEADS TO THE GENERATION AND EXTRACELLULAR RELEASE OF BETA-CLEAVED
CC SOLUBLE APP, AND A CORRESPONDING CELL-ASSOCIATED CARBOXY-TERMINAL
CC FRAGMENT WHICH IS LATER RELEASE BY GAMMA-SECRETASE (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF190727; AAF04144.1; -.
CC HSSP: P32329; 1TFS.
CC MEROPS: A01.004; -.
CC InterPro: IPR001969; Aspprotease_site.
CC InterPro: IPR001461; AspproteaseA1.
CC Pfam: PF00026; asp_1.
CC PROSITE: PS00141; ASP_PROTEASE; 1.
CC Hydrolase: Aspartyl protease; Glycoprotein; Zymogen; Transmembrane;
CC Signal.
CC SIGNAL: 1 21 POTENTIAL.
CC PROPEP 22 45 POTENTIAL.
CC CHAIN 46 501 BETA-SECRETASE.
CC DOMAIN 22 457 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 458 478 POTENTIAL.
CC DOMAIN 479 501 CYTOPLASMIC (POTENTIAL).
CC ACT_SITE 93 93 BY SIMILARITY.
CC ACT_SITE 289 289 BY SIMILARITY.
CC DISULFID 216 420 BY SIMILARITY.
CC DISULFID 278 443 BY SIMILARITY.
CC DISULFID 330 380 BY SIMILARITY.
CC CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 223 223 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 354 354 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 501 AA; 55806 MW; 24B445BC8B87DE3 CRC64;

```

Alignment Scores:

Pred. No.: 1.32e-71 Length: 501

Db	305	AlaVallySerIleLysAlaAlaSerSerThrGluLysPheProAspGlyPheIrrLeu	324
QY	1015	GGGTCGCCAGCTGGCGTGGAGCAATTCGGAACACCACTTGTCTTACTTCCTCCATAAATC	1074
Db	325	GlyGluGlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnIlePheProValIle	344
QY	1075	TCCATCTACCTGAGAGATGAGAACTCCAGAGGTANTCCGTATCACAAATCCIGCTCTAG	1134
Db	345	SerLeuTyrLeuMetGlyIleValThrAsnGlnSerPheArgIleThrIleLeuProGln	364
QY	1135	CTTTACATTCAGCCCATGATGGGGCGCGCTGAATTAAT--GAATGTTACCGATTCCGGC	1191
Db	365	GlnTyrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTyrLysPheAla	384
QY	1192	ATTTCGCCATCCCAAAATCGCTGGTATCGGTGGCCAGCGTATGGAGGGCTTCTACGTC	1251
Db	385	ValSerGlnSerSerThrGlyThrValMetGlyAlaValIleMetGluGlyPheTyrVal	404
QY	1252	ATCTTCGACAGAGCCAGAGAGGGTGGCTTCGACGAGCGCCCTCTCGACAAATTGCA	1311
Db	405	ValPheAspArgAlaAlaLysArgIleGlyPheAlaValSerAlaCysHisValHisAsp	424
QY	1312	GGTGCTGAGTGTCTGAAATTTCCGGCGCCTTTCTCAACAGAGATGTAGCCAGCACTGT	1371
Db	425	GluPheArgThrAlaAlaValGluGlyProPheValThrAlaAspMetGluAspCysGly	444
QY	1372	GTCCCGCTCAGTCTTTGACGAGCCCATTTTGTGGATTGTCTCTCTATGCGCTCATGAGC	1431
Db	445	TyrAsnIleProGlnThrAspGluSerThrLeuMetThrIleAlaTyrValMetAlaAla	464
QY	1432	GTCTGTGAGCAGCATCTCTCTTGTCTTAATCGTCTGCTGCTGCTCCCTCCGCTGTCTG	1491
Db	465	IleCys--AlaLeuPheMetLeuProLeuCysLeuMetValCysGlnTrpArgCysLeu	483
QY	1492	CGT-----CGCCCGCTGACCTGAGTCGTCGTCATGATGATCTCTCTCTG	1536
Db	484	ArgCysLeuArgHisGlnHisAspPheAlaAspPheIleSerLeuLeu	500
RESULT 3			
HACE_MOUSE			
ID	BACE_MOUSE	STANDARD;	PRT; 501 AA.
AC	P56818:		
DT	30-MAY-2000	(Rel. 39, Created)	
DI	16-OCT-2001	(Rel. 40, Last sequence update)	
DE	28-FEB-2003	(Rel. 41, Last annotation update)	
DE	Beta-secretase precursor	(EC 3.4.23.-) (Beta-site App cleaving enzyme)	
DE	(Beta-site amyloid precursor protein cleaving enzyme) (Aspartyl		
DE	protease 2) (Asp 2) (ASP2) (Membrane-associated aspartic protease 2)		
DE	(Memapsin-2).		
GN	BACE.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	MEDLINE=2002972; PubMed=10531052;		
RA	Vassar R., Bennett B.D., Babu-Khan S., Kahn S., Mendiaz E.A.,		
RA	Denis P., Teplow D.B., Ross S., Amarante P., Loeloff R., Luo Y.,		
RA	Fisher S., Fuller J., Edenson S., Lile J., Jarosinski M.A.,		
RA	Biere A.L., Curran E., Burgess T., Louis J.-C., Collins F.,		
RA	Treanor J., Rogers G., Citron M.		
RT	*Beta-secretase cleavage of Alzheimer's amyloid precursor protein by		
RT	the transmembrane aspartic protease BACE.*;		
RL	Science 286:735-741(1999).		
RN	[2]		
RN	REVIEWS TO 6 AND 81-87.		
RA	Bennett B.D., Vassar R., Citron M.		
RL	Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RP	MEDLINE=2005170; PubMed=10591213;		


```

QY 1258 GACAGAGCCAGAGAGGGTCGGCTTCGACGAGGCCCTGTGCAGAAATGCAGGTGCT 3317
RA |||||||:|||||:|||||:||||| |||
DB 407 AspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHisValHisAspGluPhe 426
QY 1318 GCAGTGTCTGAATTCGGGGCTTCTTCACAGAGGATGTAGCCAGCAACTGTGTCCCC 1377
DB |||: |||: ||||||| ||| |||: |||: |||
QY 427 ArgThrAlaAlaValGluGlyProPheValThrAlaAspMetGluAspCysGlyTyrAsn 446
DB ||| ||| ||| |||: |||: |||: |||: |||: |||: |||
QY 1378 GCTCAGTCTTTCAGCGAGGCCATTTTGATGTTGCTTCATCGCTCATGAGCGTCTGT 1437
DB ||| ||| ||| |||: |||: |||: |||: |||: |||: |||
QY 447 IleProGlnThrAspGluSerThrLeuMetThrIleAlaTyrValMetAlaAlaIleCys 466
DB |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 1438 GGAGCCATCTCTTGTCTTAATCTGCTGCTGCTGCTGCCCTCGGTGTGAGTGT--- 1494
DB |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 467 ---AlaLeuPheMetLeuProLeuCysLeuMetValCysGlnTyrArgCysLeuArgCys 485
DB |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
QY 1495 ---CCGCCCGTGCACCTGAGTGTGCTCATGATGAGTCTCTCTG 1536
DB ||| ||| ||| |||: |||: |||: |||: |||: |||: |||
QY 486 LeuArgHisGlnHisAspPheAlaAspPheAlaAspPheSerLeuLeu 500
DB ||| ||| ||| |||: |||: |||: |||: |||: |||: |||

RESULT 4
BACE_HUMAN
ID BACE_HUMAN STANDARD: PRT: 501 AA.
AC P56817; Q9BYB9; G9BYC0; Q9BYC1; Q9UJ75;
DT 30-MAY-2000 (Rel. 39, Created)
DE 30-MAY-2000 (Rel. 39, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Beta-secretase precursor (EC 3.4.23.-) (beta-site APP cleaving enzyme)
DE (Beta-site amyloid precursor protein cleaving enzyme) (Aspartyl
DE protease 2) (Asp 2) (ASP2) (Membrane-associated aspartic protease 2)
DE (Memapsin-2).
GN BACE OR BACE1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM A).
RC TISSUE=Brain;
RX MEDLINE=20002972; PubMed=10531052;
RA Vassar R., Bennett B.D., Babu-Khan S., Kahn S., Mendiola E.A.,
RA Denis P., Teplow D.B., Ross S., Amarante P., Loeloff R., Luo Y.,
RA Fisher S., Fuller J., Edenson S., Lile J., Jarosinski M.A.,
RA Biere A.L., Curran E., Burgess T., Louis J.C., Collins F.,
RA Treanor J., Rogers G., Citron M.;
RA "Beta-secretase cleavage of Alzheimer's amyloid precursor protein by
RT the transmembrane aspartic protease BACE.";
RL Science 286:735-741(1999).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM A). SEQUENCE OF 46-68, AND
RP CHARACTERIZATION.
RC TISSUE=Brain;
RX MEDLINE=20057171; PubMed=10591214;
RA Sinha S., Anderson J.P., Barbour R., Basi G.S., Caccavello R.,
RA Davis D., Doan M., Dovey H.F., Frigon N., Hong J., Jacobson-Croak K.,
RA Jewett N., Keim P., Knops J., Lieberburg I., Power M., Tan H.,
RA Tetsuno G., Tung J., Schenk D., Seubert P., Suemasa S.M., Wang S.,
RA Walker D., Zhao J., McConlogue L., Varghese J.;
RA "Purification and cloning of amyloid precursor protein beta-secretase
RT from human brain.";
RL Nature 402:537-540(1999).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM A).
RX MEDLINE=20057170; PubMed=10591213;
RA Yan R., Bienkowski M.J., Shuck M.E., Miao H., Tori M.C., Pauley A.M.,
RA Brashier J.R., Stratman N.C., Mathews W.R., Buhl A.E., Carter D.R.,
RA Tomasselli A.G., Parodi L.A., Heinrich R.L., Gurney M.E.;
RA "Membrane-anchored aspartyl protease with Alzheimer's disease beta-
RT secretase activity.";
RL Nature 402:533-537(1999).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM A).

```

```

RX MEDLINE=20120043; PubMed=10656250;
RA Hussain I., Powell D.J., Howlett D.R., Tew D.G., Meek T.D.,
RA Chapman C., Gloger I.S., Murphy K.E., Southan C.D., Ryan D.M.,
RA Smith T.S., Simmons D.L., Walsh F.S., Dingwall C., Christie G.;
RT "Identification of a novel aspartic proteinase (Asp 2) as beta-
RT secretase.";
RL Mol. Cell. Neurosci. 14:419-427(1999).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM B).
RC TISSUE=Brain, and Pancreas;
RA Michel B., De Pietri Tonelli D., Zaccchetti D., Keller P.;
RT "New beta-site APP cleaving enzyme isoform (BACE-1B) obtained from
RT human brain and pancreas.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM C).
RC TISSUE=Pancreas;
RA Zaccchetti D., De Pietri Tonelli D., Schnurbus R.;
RT "New beta-site APP cleaving enzyme isoform (BACE-1C) obtained from
RT human pancreas.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A. (ISOFORMS B; C AND D).
RC TISSUE=Brain;
RX MEDLINE=21408467; PubMed=11516562;
RA Tanahashi H., Tabira T.;
RT "Three novel alternatively spliced isoforms of the human beta-site
RT amyloid precursor protein cleaving enzyme (BACE) and their effect on
RL Neurosci. Lett. 307:9-12(2001).
RN [8]
RP SEQUENCE OF 14-501 FROM N.A. (ISOFORM A), AND CHARACTERIZATION.
RX MEDLINE=20144060; PubMed=10677483;
RA Lin X., Koelsch G., Wu S., Downs D., Dashti A., Lang J.;
RT "Human aspartic protease memapsin 2 cleaves the beta-secretase site of
RL beta-amyloid precursor protein.";
RN [9]
RP DISULFIDE BONDS.
RX MEDLINE=21950860; PubMed=11953458;
RA Fischer F., Molinari M., Bodendorf U., Paganetti P.;
RT "The disulphide bonds in the catalytic domain of BACE are critical but
RL not essential for amyloid precursor protein processing activity.";
RX J. Neurochem. 80:1079-1088(2002).
CC -!- FUNCTION: RESPONSIBLE FOR THE PROTEOLYTIC PROCESSING OF THE
CC AMYLOID PRECURSOR PROTEIN (APP). CLEAVES AT THE AMINO TERMINUS OF
CC THE A-BETA PEPTIDE SEQUENCE, BETWEEN RESIDUES 671 AND 672 OF APP.
CC LEADS TO THE GENERATION AND EXTRACELLULAR RELEASE OF BETA-CLEAVED
CC SOLUBLE APP, AND A CORRESPONDING CELL-ASSOCIATED CARBOXY-TERMINAL
CC FRAGMENT WHICH IS LATER RELEASED BY GAMMA-SECRETASE.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Name=A; Synonyms=BACE-1A, BAC-501;
CC IsoId=P56817-1; Sequence=Displayed;
CC Name=B; Synonyms=BACE-1B, BACE-I-476;
CC IsoId=P56817-2; Sequence=VSP_005223;
CC Name=C; Synonyms=BACE-1C, BACE-I-457;
CC IsoId=P56817-3; Sequence=VSP_005222;
CC Name=D; Synonyms=BACE-1D, BACE-I-432;
CC IsoId=P56817-4; Sequence=VSP_005222, VSP_005223;
CC -!- TISSUE SPECIFICITY: BRAIN.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF190725; AAF04142.1;
CC DR

```

DR EMBL; AF201468; AAF18982.1; -
 DR EMBL; AF200343; AAF17079.1; -
 DR EMBL; AF204943; AAF26367.1; -
 DR EMBL; AF338816; AAK38374.1; -
 DR EMBL; AF338817; AAK38375.1; -
 DR EMBL; AB050436; BAB40931.1; -
 DR EMBL; AB050437; BAB40932.1; -
 DR EMBL; AB050438; BAB40933.1; -
 DR EMBL; AF200193; AAF13715.1; -
 DR PIR; A59090; A59090.
 DR PDB; 1M4H; 28-AUG-02.
 DR MEROPS; A01.004; -
 DR Genew; HGNC:933; BACE.
 DR MIM; 604252; -
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.
 DR GO; GO:0008798; F:beta-aspartyl-peptidase activity; TAS.
 DR GO; GO:0009405; P:pathogenesis; TAS.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; TAS.
 DR InterPro; IPR001969; Asprotease_1.
 DR InterPro; IPR001461; Asprotease_1.
 DR Pfam; PF00026; asp; 1.
 DR PROSITE; PRO0792; PEPsin.
 DR PROSITE; PS00141; ASP_PROTEASE; 1.
 KW Hydrolase; Aspartyl protease; Glycoprotein; Zymogen; Transmembrane;
 KW Signal; Alternative splicing; 3D-structure.
 FT SIGNAL 1 21
 FT PROPEP 22 45
 FT CHAIN 46 501
 FT DOMAIN 22 457
 FT TRANSMEM 458 478
 FT DOMAIN 479 501
 FT ACT_SITE 93 93
 FT ACT_SITE 289 289
 FT DISULFID 216 420
 FT DISULFID 278 443
 FT DISULFID 330 380
 FT CARBOHYD 153 153
 FT CARBOHYD 172 172
 FT CARBOHYD 223 223
 FT CARBOHYD 354 354
 FT VARSPPLIC 146 189
 FT VARSPPLIC 190 214
 FT VARSPPLIC 214 214
 FT SEQUENCE 501 AB; 55763 MW; 377CE4C824ACEF05 CRC64;
 Alignment Scores:
 Pred. No.: 4,85e-71 Length: 501
 Score: 1178.50 Matches: 240
 Percent Similarity: 62.04% Conservative: 82
 Best Local Similarity: 46.24% Mismatches: 164
 Query Match: 35.29% Indels: 33
 Dbs: 1 Gaps: 9
 US-09-806-194-1 (1-1804) x BACE_HUMAN (1-501)
 QY 19 GCGTGTGCTGCTCTGCTGCGGCGGAGTCTGCTGCGCGCGCGCGGAGCTGCGGCC 78
 Db 2 AlaGlnAlaLeuProTrpLeuLeuLeuTrpMet-----GlyAlaGlyValLeuPro :8
 QY 79 GCGGCCCTCAGG-----CTGCCCTCGGGTGGCGGGCGGCGGCGGCGGCGGCGGCGG 123
 Db 19 AlaHisGlyThrGlnHisGlyLeuArgLeuProLeuArgSerGlyLeuGly----- 35
 QY 124 GTAGTTGCGGCCACCCCGGAGCCCGGACCCCTGCGGAGCCCGGCGGCGGCGGCGGCGGCGG 183
 Db 36 -----GlyAlaProLeu-----GlyLeuArg 42
 QY 184 CTCGCCCTGGAG-----CCTGCCCTGGCGTCCCGCGGCGGCGGCGGCGGCGGCGGCGG 237
 Db 43 LeuProArgGluThrAspGluGluProGluGluProGlyArgArgGlySerPheValGlu 62
 QY 238 ATGTAGACAACTGTCAGGGGAGTCTGCGCGGCGGCTACTACCTGGAGATGCTGATCGGG 297

Db 63 MetValAspAsnLeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetThrValGly 82
 QY 298 ACCCCCGCGCAGAGGTACAGACTTCGCTTGACACTGGAGAGAGTAACTTTGCGGTGSCA 357
 Db 83 SerProGlnThrLeuAsnLeuValAspThrGlySerSerAsnPheAlaValGly 102
 QY 358 GGAACCCCGCAGTCTCATAGACAGCTACTTTGACACAGAGAGTCTAGCACATACCOC 417
 Db 103 AlaAlaProHisProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyrArg 122
 QY 418 TCCAGGGCTTTGACCTCAGCTAGTACACAGAGAGCTGGACGGCTTCGTTGGG 477
 Db 123 AspLeuArgLysGlyValTyrValProTyrThrGlnGlyLysTrpGluGlyLeuGly 142
 QY 478 GAAGACTCGTCACCATCCCAAGGCTTCAATACTTCTTTCTTCTGCAACTTGCACCT 537
 Db 143 ThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAla 162
 QY 538 ATTTTGAATCAGAGAAATTTCTTTTTCCTGGATTAATGAATGAATGAATGAATGAAT 597
 Db 163 IleThrGluSerAspLysPhePheIleAsnGlySerAsnTrpGluGlyIleLeuGlyLeu 182
 QY 598 GCTATGTCACACTTCCCAAGCCATCAAGTCTCTGAGACCTTCTTCCGACTCCCTGGTG 657
 Db 183 AlatyAlaGluIleAlaArgProAspSerLeuGluProPhePheAspSerLeuVal 202
 QY 658 ACACAGCAACATCCCAAGCTTTCTTCCATGACAGATGTGGAGCCGCTTGCCTGT 717
 Db 203 LysGlnThrHisValProAsnLeuPheSerLeuGlnLeuCysGlyAlaGlyPheProLeu 222
 QY 718 GCTGGATCT-----GGGACCAAGGAGGTAGTCTTCTGCTGGTGAATGAACCA 768
 Db 223 AsnGlnSerGluValLeuAlaSerValGlyGlySerMetIleLeuGlyIleAspHis 242
 QY 769 AGTTTGTATAAGGACATCTTGATACCCCTATTAAAGSAAGAGTGGTACTACCAGATA 828
 Db 243 SerLeuTyrThrGlySerLeuTrpTyrThrProIleArgGluTrpTyrGluVal 262
 QY 829 GAAATCTGAATTTGAATTTGGAGCCCAAGCCTTAATCTGGAGTCGACAGATATAC 888
 Db 263 IleIleValArgValGluIleAsnGlyGlnAspLeuLysMetAspCysLysGluTyrAsn 282
 QY 889 GCAGACAAGCCACTGCTGGACAGTGGCACCACCTGCTGCGCTGCCCGAGAGGTGTTT 948
 Db 283 TyrAspLysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysValPhe 302
 QY 949 GATCGGTGGTGAAGCTCTGGCCCGCGCATCTCTGATTCAGAAATTCCTGATGTTTC 1008
 Db 303 GluAlaAlaValLysSerIleLysAlaAlaSerSerThrGluLysPheProAspGlyPhe 322
 QY 1009 TGGACTGGGTCCCGAGCTGGCGTCTGGAGCAATTCGGAACACCTTGGTCTTACTTCCCT 1068
 Db 323 TrpLeuGlyGluGlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnIlePhePro 342
 QY 1069 AAAATCTCCATCTACCTGAGAGATGAGAACTCCAGCAGGTCTTCGCTATCACAATCTTG 1128
 Db 343 ValIleSerLeuTyrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIleLeu 362
 QY 1129 CTTACGCTTTACATTCAGCCCATATGTTGGGGCGGCGCTGAATAT---GAATGTTACCGA 1185
 Db 363 ProGlnGlnTyrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTyrLys 382
 QY 1186 TTGGGCAATTTCCCATCCCAAAATGGCTGGTGAICGCTGCCAGCGTATGGAGGCTTC 1245
 Db 383 PheAlaIleSerGlnSerSerThrGlyThrValMetGlyAlaValIleMetGluGlyPhe 402
 QY 1246 TAGCTCATCTTCACAGAGCCAGAGAGGTGGGTTCGACGAGAGCCCTTCGACAGAA 1305
 Db 403 TyrValValPheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHisVal 422
 QY 1306 ATTGCAAGGTGCTGCAGTGTCTGAATTTTCGGGCGCTTCTTCAACAGAGGATGTAGCCAGC 1365

Db 207 GlnGlySerSerGlyGlyAlaValValPheGlyGlyValAspSerSerLeuTyrThrGly 226
 Qy 784 GACATCTGGTATACCCCTATTAAAGGAAGAGTGGTACTACCAATATTCGAAATTC 843
 Db 227 GlnIleTyrPheAlaProValThrGlnGluLeuTyrTrpGlnIleGlyLeuGluPhe 246
 Qy 844 GAAATGGAGGCCAAGCCCTTAATCTGGAGCTGCGAGAGATATAACGACAGACAGCCATC 933
 Db 247 LeuIleGlyGlnAlaSerGlyTyr---CysSerGlu-----GlyCysGlnAlaIle 263
 Qy 904 GTGCAGAGTGGACACGCGTCTGCGCTGCCGCCAGAGGTGTGTGATCGGTGGTGGAA 963
 Db 264 ValAspThrGlyThrSerLeuLeuThrValProGlnGlnTyrMetSerAlaLeuLeuGln 283
 Qy 964 GCTGTGGCCCGCATCTCTGATTCCAGAAATCTCTGATGCTTCTGGAGTGGGTCGCCAG 1023
 Db 284 Ala-----ThrGlyAlaGln 288
 Qy 1024 CTGGCGTGTGGAGGATTCGGAACACCTTGGTCTTACTTC----- 1065
 Db 289 -----GluAspGluTyrGlyGlnPheLeuValAsnCysAsnSer 301
 Qy 1066 -----CCTAAATCTCCATCTACCTGAGAGATCAGAACTCCAGCAGTCAATC 1113
 Db 302 IleGlnAsnLeuProThrLeuThrPheIleIle-----AsnGlyVal 315
 Qy 1114 CGTATCAATCGCTCGCTCAGCTTTACATTCACCCCATGATGGCGCCGCTGAATTA 1173
 Db 316 GluPheProLeuProSerSerTyrIle-----LeuAsnAsn 328
 Qy 1174 GAATGTTAC---CGATTGGCATTCCCA-----TCCACAAAT 1209
 Db 329 AsnGlyTyrCysThrValGlyValGluProThrTyrLeuSerAlaGlnAsnSerGlnPro 348
 Qy 1210 GCGTGTGTGTCGGTGGCCAGCGTGTATGAGGGCTTCTACGTCTTCGACAGAGCCAG 1269
 Db 349 LeuTrpIleLeuGlyAspValPheLeuArgSerTyrTyrSerValTyrAspLeuSerAsn 368
 Qy 1270 AAGAGGTGGCTTCGACGGCAGC 1293
 Db 369 AsnArgValGlyPheAlaThrAla 376
 RESULT 6
 ID PEPC_HUMAN STANDARD; PRT; 388 AA.
 AC P20142;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Gastric precursor (EC 3.4.23.3) (Pepsinogen C).
 GN PGC.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88087276; PubMed=3335549;
 RA Hayano T., Sogawa K., Ichihara Y., Fujii-Kuriyama Y., Takahashi K.;
 RT "Primary structure of human pepsinogen C gene."
 RL J. Biol. Chem. 263:1382-1385(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89079679; PubMed=2909526;
 RA Taggart R.T., Cass L.G., Mohandas T.K., Derby P., Barr P.J., Pals G.,
 RA Bell G.I.;
 RT "Human pepsinogen C (progastricsin). Isolation of cDNA clones,
 RT localization to chromosome 6, and sequence homology with pepsinogen
 RT A.;
 RL J. Biol. Chem. 264:375-379(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RT TISSUE=Placenta;

RX MEDLINE=89290840; PubMed=2567697;
 RA Pals G., Azuma T., Mohandas T.K., Bell G.I., Bacon J.,
 RA Samloff I.M., Walz D.A., Barr P.J., Taggart R.T.;
 RT "Human pepsinogen C (progastricsin) polymorphism: evidence for a
 RT single locus located at 6p21.1-pter."
 RL Genomics 4:137-148(1989).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Wong R.N.S., Tang J.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 17-101.
 RX MEDLINE=90130402; PubMed=2515193;
 RA Athauda S.B.P., Tanji M., Kageyama T., Takahashi K.;
 RT "A comparative study on the NH2-terminal amino acid sequences and
 RT some other properties of six isozymic forms of human pepsinogens and
 RT pepsins."
 RL J. Biochem. 106:920-927(1989).
 RN [6]
 RP SEQUENCE OF 17-64.
 RX MEDLINE=83079318; PubMed=6816595;
 RA Foltmann B., Jensen A.L.;
 RT "Human progastricsin. Analysis of intermediates during activation
 RT into gastricsin and determination of the amino acid sequence of the
 RT propeptide."
 RL Eur. J. Biochem. 128:63-70(1982).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (1.62 ANGSTROMS).
 RX MEDLINE=95230687; PubMed=7714902;
 RA Moore S.A., Sielecki A.R., Cherniack M.M., Tarasova N.I., James M.N.G.;
 RT "Crystal and molecular structures of human progastricsin at 1.62-A
 RT resolution."
 RL J. Mol. Biol. 247:466-485(1995).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (2.36 ANGSTROMS).
 RX MEDLINE=98069649; PubMed=9406551;
 RA Khan A.R., Cherney M.M., Tarasova N.I., James M.N.G.;
 RT "Structural characterization of activation 'intermediate 2' on the
 RT pathway to human gastricsin."
 RL Nat. Struct. Biol. 4:1010-1015(1997).
 CC -!- CATALYTIC ACTIVITY: More restricted specificity than pepsin A, but
 CC shows preferential cleavage at Tyr-Ile bonds; high activity
 CC towards hemoglobin as substrate.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M18667; AAA60062.1; ALT_INIT.
 CC EMBL; M18659; AAA60062.1; JOINED.
 CC EMBL; M18660; AAA60062.1; JOINED.
 CC EMBL; M18661; AAA60062.1; JOINED.
 CC EMBL; M18662; AAA60062.1; JOINED.
 CC EMBL; M18663; AAA60062.1; JOINED.
 CC EMBL; M18664; AAA60062.1; JOINED.
 CC EMBL; M18665; AAA60062.1; JOINED.
 CC EMBL; M18666; AAA60062.1; JOINED.
 CC EMBL; M18667; AAA60062.1; JOINED.
 CC EMBL; M23077; AAA60063.1; JOINED.
 CC EMBL; M23069; AAA60063.1; JOINED.
 CC EMBL; M23070; AAA60063.1; JOINED.
 CC EMBL; M23071; AAA60063.1; JOINED.
 CC EMBL; M23072; AAA60063.1; JOINED.
 CC EMBL; M23073; AAA60063.1; JOINED.
 CC EMBL; M23074; AAA60063.1; JOINED.
 CC EMBL; M23075; AAA60063.1; JOINED.
 CC EMBL; J04443; AAA60074.1; JOINED.
 CC EMBL; U75272; AAB18273.1; JOINED.
 CC PIR; A29937; A29937.

FT	STRAND		319	323	
FT	TURN		324	325	
FT	TURN		326	330	
FT	STRAND		331	335	
FT	HELIX		336	338	
FT	STRAND		339	343	
FT	STRAND		344	345	
FT	STRAND		347	350	
FT	TURN		351	356	
FT	STRAND		357	363	
FT	HELIX		364	368	
FT	TURN		369	370	
FT	STRAND		371	376	
FT	TURN		377	380	
FT	STRAND		381	388	
SQ	SEQUENCE	388 AA:	42426 MW:	F862DFDC1438BB92 CRC64:	

Alignment Scores:				
Pred. No.:	4,46e-16	Length:	388	
Score:	353.00	Matches:	120	
Percent Similarity:	44.79%	Conservative:	65	
Best Local Similarity:	29.06%	Mismatches:	120	
Query Match:	10.57%	Indels:	108	
DB:	1	Gaps:	21	

US-09-806-194-1 (1-1804) x PEPC_HUMAN (1-388)

QY	154	CCTGCGCAGCCGCCAGCGCGGC---TTGGCGTCCGCTGGAGCGTCGCCCTGGCGTCC	210
Db	50	ProAlaLTPlySTyArGPheGlyAspLeuSerValThrTyrgluPro-----	65
QY	211	CCGCGGGCGCGCCCAACTTCTTGCCATGGTAGACAACTGCGAGGGGACTCTGGCGCG	270
Ub	66	-----MetAlaTyrrMetAsp-----	71
QY	271	GGCTACTACCTGGAGATGCTGATGGGACC-----CGGAGAGCTACAGATTCTCGTTGAC	330
Db	72	AlaTyrrPheGlyGluIleSe:IleGlyThrProGlnAsnPheLeuValLeuPheAsp	91
QY	331	ACTGGAAGCAGTAACCTTTGCGGTG-----GCAGGAACCCCG	366
Db	92	ThrGlySerSerAsnLeuTrpValProSerValTyrcysGlnSerGluAlaCysThrSer	111
QY	367	CACCTCCTACATAGACACTCTTTGACACAGAGAGGTCTAGCACATAACGCTCCAAGGCG	426
Db	112	HisSer-----ArgPheAsnProSerGluSerSerThrTyrrSerThrAsnGly	127
QY	427	TTTGACGTCAAGTCAGTACACACAGGAGCTGGAGGCTCGTTCGGGAAGACCTC	486
Db	:28	GlnThrPheSerLeuGlnItyrGlySerGlySerLeuThrGlyPheGlyTyrraspThr	147
QY	487	GTCACCATCCCCAAGGCTTCAATACITTTTCTTGTCACCATTCGCCACTATTTTIGAA	546
Db	148	LeuthrVal-----GlnSerileGlnValProasnGlnGluPheGlyLeu	162
QY	547	TCAGAGAATTCTTTTCCCTGGG-----ATTAAATGAATGAATACATT	591
Db	163	SerGluAsn-----GluProGlyThrAsnPheValTyralaGlnPheAspGlyIleMet	180
QY	592	GGCTAGGCTTATGCCACACTTCCCAAGCCATCAAGTTCTCTGCGAGACITCTTCGACTCC	651
Db	181	GlyLeuAlaTyrrProAlaLeuSerValAspGluAlaThr-----ThraIleMetGlnGly	198
QY	652	CTGGTGACACAACCAACATCCCCAAC---GTTTTCTCCATGCAGATGTGTGGAGCGGCG	708
Db	199	MetValGlnGluGlyAlaLeuThrSerProValPheSerValTyrrLeuSerAsnGln---	217
QY	709	TTGCGCGTGTCTGGATCTGCGACCAACGAGGTAGTCTTCTGTGGTGGTGAATTGAACCA	768
Db	218	-----GlnGlySerSerGlyAlaValValPheGlyGlyValAspSer	232
QY	769	ACTTTGTATAAGGAGACATCTGGTATACCCCTATTAAAGGAAGAGTGGTACTACCAATA	828
Db	233	SerLeuTyrrThrGlyGlnIleTyrrIleAlaProValThrGlnGluLeuTyrrtpGlnIle	252


```
QY 832 ATCTCGAAATTCGAATTCGAGCCCAACCCCTTAATCTGGACGTGAGACAGATATACCCA 891
Db 254 IleGluGluPheLeuIleGlyGlyGlnAlaSerGlyTrp---CysSerGlu-----Gly 270
QY 892 GACAAGGCGATCGTGAGAGTGGACACCGCTGCTGCGCCCTGCCAGAAAGGIGTTGAT 951
Db 271 CysGlnAlaIleValAspThrGlyThrSerLeuLeuThrValProGlnGlnTyrMetSer 290
QY 952 GCGGTGGTGGAGCTGTGGCCCGCGCATCTCTGTATCCAGAAATTCCTGTGATGGTTCTGG 1011
Db 291 AlaPheLeuGluAla----- 295
QY 1012 ACTGGGTGCCACTGGCGGTGGAGCAATTCGGAAACACCTTGTCTTACTGTC----- 1065
Db 296 ThrGlyAlaGln-----GluAspGluTyrGlyGlnPheLeuVal 308
QY 1066 -----CCTAAATCTCCATCTACCTGAGAGATGAGAACTCC 1101
Db 309 AsnCysAspSerIleGlnAsnLeuProThrLeuThrPheIleIle----- 323
QY 1102 ACCAGTCAATCCGTATCATCAATCTGCTCAGCTTTATCATCACTACCCCATGATGGGCC 1161
Db 324 ---AsnGlyValGluPheProLeuProProSerSerTyrIle----- 336
QY 1162 GGCCTGAATATGAATGTTCAC---CGATTGGCATTTCGCCA----- 1200
Db 337 ---LeuSerAsnAsnGlyTyrCysThrValGlyValGluProThrTyrLeuSerSerGln 355
QY 1201 ---TCCAAATATGCGTGTGTATCGGTGGCCAGGTGATGGAGGGTCTTACGTCACTTC 1257
Db 356 AsnSerGlnProLeuTyrPheLeuGlyValPheLeuArgSerTyrIleSerValPhe 375
QY 1258 GACAGAGCCCAAGAGGGTGGCTTCGACGAGC 1293
Db 376 AspLeuGlyAsnAsnArgValGlyPheAlaThrAla 387

RESULT 8
PEPC_CAVPO
ID PEPC_CAVPO STANDARD; PRT: 394 AA.
AC O64411;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Gastricin precursor (EC 3.4.23.3) (Pepsinogen C).
GN PGC.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92355614; PubMed=1644829;
RA Kageyama T., Ichinose M., Tsukada S., Miki K., Kurokawa K., Koiwai O.,
RA Tani M., Yakabe E., Athauda S.B., Takahashi K.;
RT "Gastric procathepsin E and progastricsin from guinea pig.
RT Purification, molecular cloning of cDNAs, and characterization of
RT enzymatic properties, with special reference to procathepsin E.";
RL J. Biol. Chem. 267:16450-16459(1992).
CC -!- CATALYTIC ACTIVITY: More restricted specificity than pepsin A, but
CC shows preferential cleavage at Tyr-Xaa bonds; high activity
CC towards hemoglobin as substrate.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL: M88652; AAA37053.1; -.
```

```
PIR: B43356; B43356.
HSSP: P20142; LAVF.
DR MEROPS; A01.003; -.
DR InterPro; IPR001969; Aspprotease_site.
DR Pfam; PF00026; asp; 1.
DR PRINTS; PR00792; PEPSIN.
DR PROSITE; PS00141; ASP_PROTEASE; 2.
KW Hydrolase; Aspartyl protease; Zymogen; Digestion; Signal.
FT SIGNAL 1 16
FT PROPEP 17 65
FT CHAIN 66 394
FT ACT_SITE 97 97
FT ACT_SITE 283 283
FT DISULFID 110 115
FT DISULFID 273 277
FT DISULFID 316 349
SQ SEQUENCE 394 AA; 42995 MW; 114F08E105D49865 CRC64;

Alignment Scores:
Pred. No.: 3,54e-14 Length: 394
Score: 324.50 Matches: 107
Percent Similarity: 46.07% Conservative: 63
Best Local Similarity: 29.00% Mismatches: 116
Query Match: 9.72% Indels: 83
DB: Gaps: 18

US-09-806-194-1 (1-1804) x PEPC_CAVPO (1-394)
QY 274 TACTACCTGGAGATGCTGTATCGGAGCCCGCCGAGAGCTACAGATTCGTGTGACACT 333
Db 79 TyrPheGlyGlnIleSerLeuGlyThrProGlnSerPheGlnValLeuPheAspThr 98
QY 334 GGAAGCAGTAACTTT-----GCCGTGGCAGCAACCCGCCAC 369
Db 99 GlySerSerAsnLeuTyrValProSerValTyrCysSerSerLeuAlaCysThrThrHis 118
QY 370 TCCTACATAGACAGCTACTTTTGACACAGAGAGGTGTAGCACATACCCGCTCCAGGGGCTT 429
Db 119 -----ThrArgPheAsnProArgAspSerSerThrTyrValAlaThrAspGln 134
QY 430 GACGTCACAGTGAAGTACACACAGAGAGCTGGACGGCTTCGTGGGAGACCTCGTC 489
Eb 135 SerPheSerLeuGluTyrGlyThrGlySerLeuThrGlyValPheGlyTyrAspThrMet 154
QY 490 ACCATC-----CCCAA---GGCTTCAATCTCTTTTCTTGCAACATT 531
Db 155 ThrIleGlnAspIleGlnValProLysGlnGluPheGlyLeuSer----- 169
QY 532 GCCACTATTTTGAATCAGAGAAATTTCTTTTTCCTGGG-----ATTAAA 576
Db 170 -----GluThrGlu-----ProGlySerAspPheValTyrAlaGlu 181
QY 577 TGGAAATGGAATACTTGGCTAGCTATATCCCACTTCCCAAGCCATCAAGTCTCTGGAG 636
Db 182 PheAspGlyIleLeuGlyLeuGlyTyrProGlyLeuSerGlyGlyAlaThr----- 199
QY 637 ACCTTCATCGACTCCCTGGTGACACAAAGCAAAATC---CCCAACGTTTCTCCATGACAG 693
Db 200 ThrAlaMetGlnGlyLeuLeuArgGluGlyAlaLeuSerGlnSerLeuPheSerValTyr 219
QY 694 ATGTGTGGAGCCGCTTGGCCGTTGCTGGATCT-----GGGACCAACGAGGAGTAGTCTT 747
Db 220 Leu-----GlySerGlnGlnGlySerAspGluGlyGlnLeu 231
QY 748 GTCTTGGTGGAAATGAAACCAAGTTTGTATTAAGGACACATCTGGTATACCCCTATTAAG 807
Db 232 IleLeuGlyGlyValAspGluSerLeuTyrThrGlyAspIleTyrThrProValThr 251
QY 808 GAACAGTGGTACTACCAAGATAGAAATCTGAAATTTGCAAAATTTGAGCCCAAGACCTTAAT 867
Db 252 GlnGluLeuTyrTrpGlnIleGlyIleGluGlyPheLeuIleAspGlySerAlaSerGly 271
```


CC -1- CATALYTIC ACTIVITY: Specificity similar to, but narrower than,
 CC that of pepsin A. Does not cleave the 4-Gln-His-5 bond in: B
 CC
 CC -1- ENZYME REGULATION: Inhibited by pepstatin.
 CC
 CC -1- SUBUNIT: Monomer.
 CC
 CC -1- SUBCELLULAR LOCATION: Lysosomal.
 CC
 CC -1- MISCELLANEOUS: The isoelectric point is 6.8. Has optimal activity
 CC at pH 2.5 with hemoglobin as the substrate and the optimal
 CC temperature is 37 degrees Celsius.
 CC
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF312364; AAC27733.1; -.
 CC
 CC HSP: P07339; ILYB.
 CC
 CC MEROPS: A01.009; -.
 CC
 CC InterPro: IPR001969; Asparticase_site.
 CC
 CC InterPro: IPR001461; AsparticaseA1.
 CC
 CC Pfam: PF00026; asp; 1.
 CC
 CC PRINTS: PR00792; PEPSIN.
 CC
 CC PROSITE: PS00141; ASP_PROTEASE; 2.
 CC
 CC Hydrolase; Aspartyl protease; Glycoprotein; Lysosome; Signal; Zymogen.
 KW
 FT SIGNAL 1 18 POTENTIAL.
 FT PROPEP 19 61 ACTIVATION PEPTIDE.
 FT CHAIN 62 396 CATHESPIN D.
 FT ACT_SITE 94 94 BY SIMILARITY.
 FT ACT_SITE 281 281 BY SIMILARITY.
 FT DISULFID 107 114 BY SIMILARITY.
 FT DISULFID 272 276 BY SIMILARITY.
 FT DISULFID 315 352 BY SIMILARITY.
 FT CARBOHYD 131 131 N-LINKED (GLCNAC...) (POTENTIAL).
 FT SEQUENCE 396 AA; 43315 MW; D0375DC38567A31B CRC64;

Alignment Scores:

Pred. No.: 1.41e-13 Length: 396
 Score: 315.50 Matches: 112
 Percent Similarity: 42.86% Conservative: 65
 Best Local Similarity: 27.12% Mismatches: 141
 Query Match: 9.45% Indels: 95
 DB: 1 Gaps: 18

US-09-806-194-1 (1-1804) x CATD_CLUHA (1-396)

Qy 148 GGGACCCCTCGGAGCGCCAGCGGCTTGGCGCTCGCTGGAGCGCTCGCTGGCG 207
 Db 47 GlyThrAsnSerLeuGlnHisAsnGlnGlyPheProSerSerAsnAlaPro----- 63
 Qy 208 TCCCGCGCGGCGCGCCAACTCTTGGCATGGTAGAACCTGCAGAGGAGACTCTGGC 267
 Db 64 ThrProGlnThrLeuLysAsnTyrMet-----Asp 73
 Qy 268 CGCGGCTACTACCTGGAGATGCTGATCGGACCCCGCCAGAGCTACAGATCTCGCT 327
 Db 74 AlaGlnTyrTyrGlyGluIleGlyLeuGlyThrProValGlnMetPheThrValValPhe 93
 Qy 328 GACACTGGAAGCAGTAACCTTTGGCTGGCAGGACCCCGCCACTCCTACATAGACACG-- 384
 Db 94 AspThrGlySerSerAsnLeuThrProSerIleHisCysSerPheThrAspIleAla 113
 Qy 385 -----TACTTTGACAGAGAGGTCTTAGCACATACCCCTCCAGGGCTTT 429
 Db 114 CysLeuLeuHisLysTyrAsnGlyAlaLysSerSerThrTyrValLysAsnGlyThr 133
 Qy 430 GACGTCACGTGAGTACACAGAGGAGCTGGACGGGCTTCGTTGGGGAGACCTCGTC 489
 Db 134 GluPheAlaIleGlnTyrGlySerGlySerLeuSerGlyTyrLeuSerGlnAspSerCys 153

Qy 490 ACCATCCCAAGGCTTCAACTCTTTTGTCAACATTGCCACACTATTTTGAATCA 549
 Db 154 ThrIle-----GlyAspIleValValGluLysGlnLeuPhe---Gly 166
 Qy 550 GAGAATTTCTTTTGGCTGGGATT-----AAATGGAATGAATACTTGGC 594
 Db 167 GluAlaIleLysGlnProGlyValAlaPheIleAlaAlaLysPheAspGlyIleLeuGly 166
 Qy 595 CTAGCTTATGCCACACITGCCAAGCCATCAAGTCTCTCGAGACCTTCTCGACTCCCTG 654
 Db 187 MetAlaTyrProArgIleSer-----ValAspGlyValProValPheAspMetMet 204
 Qy 655 GTGACACAAGCAACAATC---CCCAAGCTTTTCTCCATSCAGATGTGTGGAGCGGCTTG 711
 Db 205 MetSerGlnLysValGluGlnAsnValPheSerPheTyrLeu----- 219
 Qy 712 CCGCTTGGTGGATCTGGGACCAAC-----GGAGGTAGTCTTGTCTTGGT 756
 Db 220 -----AsnArgAsuProAspThrGluProGlyGlyGluLeuLeuGly 234
 Qy 757 GGAATTGAACCAAGTTGTATAAAGAGGACATCTGGTATACCCCTATTAAAGAGAGTGG 816
 Db 235 GlyThrAspProLysTyrTyrThrGlyAspPheAsnTyrValProValThrArgGlnAla 254
 Qy 817 TACTACCATAGAAATCTGAAATTTGGAAATTTGGAGGCCAAAGCCTTAATCTGGACTGC 876
 Db 255 TyrTrpGlnIleHisMetAspGlyMetSerIleGlySerGln---LeuThrLeu---Cys 272
 Qy 877 AGAGAGTATACGACAGCAAGCCATCTGTGGACAGTGGCCACGCTGCTGGCGCTGCC 936
 Db 273 LysAsp-----GlyCysGluAlaIleValAspThrGlyThrSerLeuIleThrGlyPro 290
 Qy 937 CAGAAGGTGTTGATCGGCTGGTGGAGCTGGCCCGCGCATCTCIGATT----- 987
 Db 291 ProAlaGluValArgAlaLeuGlnLysAlaIleGlyAlaIleProLeuIleGlnGlyGlu 310
 Qy 988 -----CCAGAAATCTCTCATGTTCTGGACT 1014
 Db 311 TyrMetIleAspCysLysLysValProThrLeuProThrIleSer-----PheAsnVal 328
 Qy 1015 GGTGCCCACTGGCGTGGAGCAATTCGAAACACCTTGGTCTTACTTCCCTAAATC 1074
 Db 329 GlyGlyLys-----ThrTyrSerLeuThrGlyGluGlnTyr----- 340
 Qy 1075 TCCATCTACTGAGATCAGAACTCCAGCAGTCAATCCGTATCATCAATCTCGCTCAG 1134
 Db 341 -----ValLeuLysGluSerGlnGlyLysThrIleCysLeuSerGlyLeuMetGly 358
 Qy 1135 CTTTACATTCAGCCCATGATGGGGCGCGCTGAATTATGAATGTACCGATTTCGCATT 1194
 Db 359 LeuGluIleProPro----- 363
 Qy 1195 TCCCATTCACAAATCCGCTGGTGTGTCGCGCAGCGTATGAGGGCTTCTACGTCATC 1254
 Db 364 ---ProAlaGlyProLeuThrIleLeuGlyAspValPheIleGlyGlnTyrTyrThrVal 382
 Qy 1255 TTCGACAGAGCCCAAGAGGCTGGCGCTTCGCGAGGAGC 1293
 Db 383 PheAspArgGluSerAsnArgValGlyPheAlaLysSer 395
 RESULT 11
 ID1_ORYSA
 ID1_ORYSA STANDARD: PRT; 509 AA.
 AC Q42456;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Aspartic proteinase oryzasin 1 precursor (EC 3.4.23.-).
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OC NCBI_TaxID=4530;

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; D00215; BAA00153.1; .
DR PIR; A41443; A41443.
DR HSSP; P00794; 4CMS.
DR MEROPS; A01.028; .
DR InterPro; IPR001969; Aspprotease.site.
DR InterPro; IPR001461; AspproteaseA1.
DR Pfam; PF00026; asp; 1.
DR PRINTS; PRO0792; PEPSIN.
DR PROSITE; PS00141; ASP_PROTEASE; 2.
KW Hydrolase; Aspartyl protease; Digestion; Signal; Glycoprotein.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 383 EMBRYONIC PEPSINOGEN.
FT ACT_SITE 94 94 BY SIMILARITY.
FT ACT_SITE 276 276 BY SIMILARITY.
FT DISULFID 107 112 BY SIMILARITY.
FT DISULFID 267 271 BY SIMILARITY.
FT DISULFID 310 344 BY SIMILARITY.
FT CARBOHYD 132 132 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 204 204 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 309 309 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 350 350 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARIANT 51 51 T -> S.
SQ SEQUENCE 383 AA; 41719 MW; 1642796871611F54 CRC64;

Alignment Scores:
Pred. No.: 3,26e-13 Length: 383
Score: 310.00 Matches: 106
Percent Similarity: 42.78% Conservative: 63
Best Local Similarity: 26.84% Mismatches: 136
Query Match: 9.28% Indels: 90
DB: 1 Gaps: 15

US-09-806-194-1 (1-1804) x PEPE_CHECK (1-383)

QY 166 CACGCC-----GAGGGTTGGCGTCGGCCCTGGAGCCTGGCCCTGCCCTGCCCGCGGCG 219
DB 55 HisAlaPheProAspValLeuThrValValThrGluProLeuLeu----- 69
QY 220 GCCGCCAAGCTTCTGGCCATGTGTAGACACACCTGCGAGGGGACTCTGGCGGGGCTACTAC 279
DB 70 -----AsnThrLeuAspMet-----GluTyrTyr 77
QY 280 CTGGAGATGCTGATCGGAGCCCGCCGAGAGGTACAGATTCTGTTGACACTGGAAGC 339
DB 78 GlyThrIleSerIleGlyThrProGlnAspPheThrValValPheAspThrGlySer 97
QY 340 AGTAAGTTGGCGTGGCAGGA-----ACCCGCGCACTCTACATACACACGATAC 387
DB 98 SerAsnLeuTrpValProSerValSerCysThrSerProAlaCysGlnSerHisGlnMet 117
QY 388 TTGACACAGAGGCTTAGCACATACCCCTCCAGGCGCTTGGAGCTCAGCTCAGCTGAAGTAC 447
DB 118 PheAsnProSerGlnSerSerThrTyrLysSerThrGlyGlnAsnLeuSerIleHisTyr 137
QY 448 ACACAGAGGAGCTGGAGGGCTTGGGGAAGACCTCGTCACCATCCCAAGGCTTC 507
DB 138 GlyThrGlyAspMetGluGlyThrValGlyCysAspThrValThrValAlaSerLeuMet 157
QY 508 AATAGCTTTTCTTGTCAACATGCCACTATTTTGAATCAGAGAATTTCTTTTGGCT 567
DB 158 AspThrAsnGlnLeuPheGlyLeuSerThr-----SerGluProGlyGlnPheVal--- 175
QY 568 GGGATTAATGAATGGATATCTGGCTAGCTTAAGCCACACTTGCACAGCCATCAAGT 627
DB 176 TyrValLysPheAspGlyIleLeuGlyTyrProSerLeuAla-----AlaAsp 193
QY 628 TCTCTGGAGACCTTCTCGACTCCCTGGTGCACAGCAAGCAAAACATC---CCCAAGCTTTTC 684
DB 194 GlyIleThrProValPheAspAsnMetValAsnGluSerLeuLeuGluGlnAsnLeuPhe 213

655 GTGACACAGCAAAACATCCCAAGCTTTCTCCATGCAGATGTCGAGCGGCTGGCC 714
203 LeuGlyGlu-----GlyAlaLeuSerGlnProLeuPheGlyValTyrLeu--- 217
715 GTTGTCTGATCT-----GGACCAAGGAGGTAGTCTTGTCTTGGTGGAAATGAACCA 766
218 -----GlySerGlnGlnGlySerAsnGlyGlnIleValPheGlyGlyValAspLys 235
769 AGTTTGTATAAGGAGACATCTGGTATACCCCTATTAAAGAAAGTGGTACTACCAAGATA 828
236 AsnLeuTyrThrGlyGluIleThrTrpValProValThrGlnGluLeuTyrTrpGluIle 255
829 GAAATTCGAAATGGAAATGGAGCGCAAGGCTTAATCTGACTGACAGAGTATAAC 888
256 ThrIleAspPheLeuIleGlyAspGlnAlaSerGlyTrp---CysSerSerGlnGly 274
889 GCAGACAGGCCATCTGGGACAGTGGCAGCCAGCAGCTGTGGCGCTGCCCGCCAGAGGTGTTT 948
275 Cys---GlnGlyIleValAspThrGlyThrSerLeuLeuValMetProAlaGlnTyrLeu 293
949 GATCGGTGTGGAGCTGTGGCCCGCGCATCTCTGATTCACAGATCTCTGATGGTTTC 1008
294 SerGluLeuLeuGlnThrIleGlyAlaGlnGlu-----GlyGluTyrGlyGluTyrPhe 311
1009 TGGACTGGTCCCGACCTGGCTGTGTCAGCAATTCGGAAACACCTTGTCTTACTTCCTT 1068
312 -----ValSerCysAspSerValSerSer-----LeuPro 321
1069 AAATCTCCATCTACCTGAGAGATGAGAACTCCAGCAGGTCTTCCTGATCACAATCTG 1126
322 ThrLeuSerPheValLeu-----AsnGlyValGlnPheProLeuSer 335
1129 CCTCAGCTTTAC-----ATTACGCCCATGATCGGGCGCCGCTGAATTAAGTAATCTACCGA 1185
336 ProSerSerTyrIleIleGlnGluAspAsnPheCysMetValGlyLeuGluSerIleSer 355
1186 TTGGCATTTCCCATCCCAAAATGCGCTGTGATCGTGGTGGTGGTGGTGGTGGTGGTGGT 1245
356 LeuThrSerGluSerGlyGlnProLeuTrpIleLeuGlyAspValPheLeuA-gSerTyr 375
1246 TACGTCATCTCGACAGAGCCCGAGAGAGGTGGCTTCGCAGCAGC 1293
376 TyrAlaIlePheAspMetGlyAsnAsnLysValGlyLeuAlaThrSer 391

RESULT 13
ID PEPE_CHECK STANDARD: PRT; 383 AA.
AC P16476;
DI 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
OS Embryonic pepsinogen precursor (EC 3.4.23.-).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phas. gallinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98227903; PubMed=3131317;
RA Hayashi K., Agata K., Mochii M., Yasugi S., Eguchi G., Mizuno T.;
RT "Molecular cloning and the nucleotide sequence of cDNA for embryonic
RT chicken pepsinogen: phylogenetic relationship with prothymosin.";
RL J. Biochem. 103:290-296(1988).
CC -!- DEVELOPMENTAL STAGE: SPECIFICALLY SECRETED DURING THE EMBRYONIC
CC PERIOD IN THE CHICKEN PROVENTRICULUS (GLANDULAR STOMACH).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

QY 685 TCATCCACATGTCGGACCCCGCTTCCGCGTTCCTCGATCTCGACCAACCGGACGAGT 744
DB |||||:::
DB 214 SerValTyrLeuSerArgGluProMet:::-----GlySerMet 225
QY 745 CTTGCTCTGGTGGGAATTGAACCAAGTTTGATATAAGGACAGACATCGTGTATACCCATT 804
DB |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 226 ValValPheGlyGlyIleAspGluSerTyrPheThrGlySerIleAsnTrpIleProVal 245
QY 805 AAGGAAGAGTGTACTACACAGATGAATCTGAAATTTGGAAGTGGAGGCAAGAGCTT 864
DB |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 246 SerTyrGlnGlyTyrTrpGlnIleSerMetAspSerIleValAsnLysGlnGluIle 265
QY 865 AATCTGCAGTCGACAGACATATAACCCACAGACCCATCGTCGACAGTCCGACCGCTG 924
DB |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB 266 AlaCysSer:::-----SerGlyCysGlnAlaIleIleAspThrGlyThrSerLeu 281
QY 925 CTGCGCCCTCCGCCAGAGGTGTTTCATCGGTGGTGAAGCTGTGGCCCGGCATCTCTG 984
DB |||||:::|||||:::|||||:::|||||:::|||||:::
DB 282 ValAlaGlyProAlaSerAspIleAsnAspIleGlnSerAlaValGly:::----- 297
QY 985 ATTCAGAAATCTCTGATGGTTCCTGGACTGGTCCGACGCTGGTGGACGAAATTCG 1044
DB |||||:::|||||:::|||||:::|||||:::|||||:::
DB 298 :::-----AlaAsnGln 303
QY 1045 GAACACCTTGCTTACTTCCCTAAATCTCCATCTACCTCAGAGATGAGAACTCCAGC 1104
DB |||||:::|||||:::|||||:::|||||:::
DB 301 AsnThrTyrGlyGluTyr:::-----SerVal 309
QY 1105 AGGTCATTCGGTATCACAACTCGCTCAGCTTACATTCAGCCCATGATGGGSCGCGC 1164
DB |||||:::|||||:::|||||:::|||||:::
DB 309 AsnCysSerHisIleLeuAlaMetProAspValValPhe:::-----ValIleGly--Gly 325
QY 1165 CTGAATTAT:::-----GATGTCATC 1182
DB |||||:::
DB 326 IleGlnTyrProVal::ProAlaLeuAlaIyrThrGluGlnAsnGlyGlnGlyThrCysMet 345
QY 1183 CGATTCGGCATTTCCCGCCACCAATCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1242
DB |||||:::|||||:::|||||:::|||||:::
DB 346 SerSerPheGlnAsnSerSerAlaAsnLeuTrpIleLeuGlyAspValPheIleArgVal 365
QY 1243 TTCTACGTCATCTTCGACAGCCGACAGAGAGGGTGGGCTTCGCA 1287
DB |||||:::|||||:::|||||:::|||||:::
DB 366 TyrTyrSerIlePheAspArgAlaAsnAsnArgValGlyLeuAla 380
RESULT 14
CATD_HUMAN
ID CATD_HUMAN STANDARD; PRT: 412 AA.
AC P07339;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Cathepsin D precursor (EC 3.4.23.5).
GN CTSD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85270436; PubMed=3527292;
RA Faust P.L., Kornfeld S., Chirgwin J.M.;
RT "Cloning and sequence analysis of cDNA for human cathepsin D.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:4910-4914(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87231068; PubMed=3588310;
RA Westley B.R., May F.E.B.;
RT "Oestrogen regulates cathepsin D mRNA levels in oestrogen responsive
human breast cancer cells.";
RL Nucleic Acids Res. 15:3773-3786(1987).
RN [3]
RP SEQUENCE FROM N.A.

```

```

RX
RA MEDLINE=91299158; PubMed=2069717;
RA Redeker B., Heckendorf B., Grosch H.W., Mersmann G., Hasilik A.;
RT "Molecular organization of the human cathepsin D gene.";
RL DNA Cell Biol. 10:423-431(1991).
RN [4]
RP SEQUENCE FROM N.A.
RX TISSUE=Kidney;
MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP SEQUENCE OF 1-22 FROM N.A.
MEDLINE=94085791; PubMed=8262386;
RA May F.E., Smith D.J., Westley B.R.;
RT "The human cathepsin D-encoding gene is transcribed from an estrogen-
regulated and a constitutive start point.";
RL Gene 134:277-282(1993).
RN [6]
RP SEQUENCE OF 1-22 FROM N.A.
MEDLINE=95021303; PubMed=7935485;
RA Augereau P., Miralles F., Cavailles V., Gaudelet C., Parker M.,
RA Rochefort H.;
RT "Characterization of the proximal estrogen-responsive element of
human cathepsin D gene.";
RL Mol. Endocrinol. 8:693-703(1994).
RN [7]
RP SEQUENCE OF 170-180.
RA I-SSUE=Liver;
RA Hochstrasser D.F., Frutiger S., Paquet N., Bairoch A., Ravier F.,
RA Pasquali C., Sanchez J.-C., Tissot J.-D., Bjellqvist B., Vargas R.,
RA Appel R.D., Hughes G.J.;
RT Submitted (JUN-1992) to the SWISS-PROT data bank.
RN [8]
RP VARIANT VAL-58.
MEDLINE=20179010; PubMed=10716266;
RA Papasotiropoulos A., Bagli M., Kurz A., Kornhuber J., Forstl H.,
RA Maier W., Pauls J., Lautenschlager N., Heun R.;
RT "A genetic variation of cathepsin D is a major risk factor for
Alzheimer's disease.";
RL Ann. Neurol. 47:399-403(2000).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
RX TISSUE=Spleen;
MEDLINE=93223670; PubMed=8467789;
RA Metcalf P., Fusek M.;
RT "Two crystal structures for cathepsin D: the lysosomal targeting
signal and active site.";
RL EMBO J. 12:1293-1302(1993).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX TISSUE=Liver;
MEDLINE=93342076; PubMed=8393577;
RA Baldwin E.T., Bhat T.N., Gulnik S., Hosur M.V., Sowder R.C. II,
RA Cachau R.E., Collins J., Silva A.M., Erickson J.W.;
RT "Crystal structures of native and inhibited forms of human cathepsin

```

D: implications for lysosomal targeting and drug design.";
 Proc. Natl. Acad. Sci. U.S.A. 90:6796-6800(1993).
 -I- FUNCTION: Acid protease active in intracellular protein breakdown.
 Involved in the pathogenesis of several diseases such as breast
 cancer and possibly Alzheimer's disease.
 -I- CATALYTIC ACTIVITY: Specificity similar to, but narrower than,
 that of pepsin A. Does not cleave the 4-Gln-I-His-5 bond in B
 chain of insulin.
 -I- SUBUNIT: CONSISTS OF A LIGHT CHAIN AND A HEAVY CHAIN.
 -I- SUBCELLULAR LOCATION: Lysosomal.
 -I- POLYMORPHISM: The Val-58 allele is significantly overrepresented
 in demented patients (11.8%) compared with nondemented controls
 (4.9%). Carriers of the Val-58 allele have a 3.1-fold increased
 risk for developing AD than noncarriers.
 -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL: M11233; AAB59529.1; -
 EMBL: X05344; CAA28955.1; -
 EMBL: M63138; AAA51922.1; -
 EMBL: M63134; AAA51922.1; JOINED.
 EMBL: M63135; AAA51922.1; JOINED.
 EMBL: M63136; AAA51922.1; JOINED.
 EMBL: M63137; AAA51922.1; JOINED.
 EMBL: BC016320; AAI16320.1; -
 EMBL: L12980; AAA16314.1; -
 EMBL: S74689; AAD14156.1; -
 EMBL: S52557; AAD13868.1; -
 F1R: A25771; KHRUD.
 PDB: 1LYA; 31-JAN-94.
 PDB: 1LYB; 31-JAN-94.
 PDB: 1LYW; 22-JUL-99.
 MEROPS: A01.009; -
 SWISS-2DPAGE: P07339; HUMAN.
 SlenA-2DPAGE: P07339; -
 GeneW: HGNC:2529; CTSD.
 MIM: 116840; -
 GO: GO:0004192; F:cathepsin D activity; TAS.
 InterPro: IPR001969; Asparticase.site.
 InterPro: IPR001461; Asparticase.
 Pfam: PF00026; asp.1.
 PRINTS: PR00792; PEPsin.
 PROSITE: PS00141; ASP_PROTEASE; 2.
 Hydrolase: Aspartyl protease; Glycoprotein; Lysosome; Signal; Zymogen;
 Polymorphism: Alzheimer's disease; 3D-structure.
 SIGNAL 1 18
 PROPEP 19 64
 CHAIN 65 412
 CHAIN 65 161
 CHAIN 169 412
 ACT_SITE 97 97
 ACT_SITE 295 295
 DISULFID 91 160
 DISULFID 110 117
 DISULFID 286 290
 DISULFID 329 366
 CARBOHYD 134 134
 VARIANT 263 263
 N-LINKED (GLCNAC. . .).
 N-LINKED (GLCNAC. . .).
 A -> V (ASSOCIATED WITH INCREASED RISK IN
 AD; POSSIBLY INFLUENCES SECRETION AND
 INTRACELLULAR MATURATION; dbSNP:17571).
 /FTid=VAR_011621.
 STRAND 67 74
 TURN 75 77
 STRAND 78 85
 TURN 86 89

FT	STRAND	90	97
FT	TURN	98	99
FT	STRAND	103	107
FT	TURN	108	109
FT	TURN	112	113
FT	HELIX	115	118
FT	TURN	119	123
FT	STRAND	123	123
FT	HELIX	123	127
FT	TURN	129	130
FT	STRAND	132	141
FT	STRAND	146	158
FT	STRAND	172	184
FT	HELIX	188	192
FT	STRAND	197	200
FT	HELIX	204	206
FT	STRAND	208	210
FT	HELIX	214	220
FT	TURN	221	222
FT	STRAND	228	233
FT	STRAND	243	247
FT	TURN	248	254
FT	HELIX	252	254
FT	STRAND	255	263
FT	STRAND	267	267
FT	TURN	268	269
FT	STRAND	270	279
FT	TURN	280	281

Alignment Scores: 4.16e-13 Length: 412
 Pred. No.: 308.50 Matches: 121
 Score: 43.85% Conservative: 75
 Percent Similarity: 27.07% Mismatches: 180
 Best Local Similarity: 9.24% Indels: 71
 Query Match: 1 Gaps: 22
 DB: 1

US-09-806-194-1 (1-1804) x CATD_HUMAN (1-412)
 QY 25 CTGCTGCTCTGCTGCGCCAGTGGCTCTGCGCGCGCGCGCGAGCTGGCCCGCGGCC 84
 DB 6 LeuLeuProLeuAlaLeu-----CysLeuLeuAlaPro-----AlaSerAlaLeu 21
 QY 85 TTCACGCTGCCCTCGGGTGGCGCGCGCGCGAGCGGTAGTTGCGCCCGCGCGGGA 144
 DB 22 ValArgIleProLeuHisLysPheThrSerIleArgArgThrMetSerGluValGlyGly 41
 QY 145 -----CCCGGAGCCCTGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTTC 186
 DB 42 SerValGluAspLeuIleAlaLysGlyProValSerLysTyrSerGlnAlaValProAla 61
 QY 187 GCCTTGAGCGCTGCCCTGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTTC 246
 DB 62 ValThrGluGlyProIle-----ProGluValLeuLysAsnTyrMet----- 75
 QY 247 AACCTGCGAGGGGACTCTGCGCGCGCGCTACTACCTGGAGATGCTGATCGGACCGCCGCG 306
 DB 76 -----AspAlaGlnTyrTyrGlyGluIleGlyIleGlyThrProPro 89
 QY 307 CAGAGCTACAGATCTCGTTGACACTGGAAGCAGTACTTTGCGCGTGGCAGACCGCGC 366
 DB 90 GlnCysPheThrValValPheAspThrGlySerSerAsnLeuTrpValProSerIleHis 109
 QY 367 CACTCTCATATAGACAGC-----TACTTTGACACAGAGAGGTCTAGC 408
 DB 110 CysLysLeuLeuAspIleAlaCysTrpIleHisLysLysTyrAsnSerAspLysSer 129
 QY 409 ACATACCGCTCCAAAGGGCTTTGACGCTACAGTACAGTACACAGAGAGTGGACGGCGC 468
 DB 130 ThrTyrValLysAsnGlyThrSerPheAspIleHisTyrGlySerLysLeuSerGly 149
 QY 469 TTCGTTGGGGAAGACCTCGTCCACCATCCCG-----AAAGCGCTCAATACTTCTTTCTT 522
 DB 522

```

Db      150  TyrLeuserGlnAspThrValSerValProCysGlnSerAlaSerSerAlaSerAlaLeu 169
QY      523  -----GTCAACATGCCACTATTTTGTGAATCAGAGAAATTTCTTTGCTCGGATT--- 573
Db      170  GlyGlyVallyValGluArgGlnValPheGlyGluAlaThrLysGlnProGlyLeu 189
QY      574  -----AAATGGAAATGGAATGCTTGGCTAGCTTATGCCACACTGCCAAGCCA 621
Db      190  PheIleAlaAlaLysPheAspGlyIleLeuGlyMetAlaTyrProArgIleSer----- 207
QY      622  TCAAGTCTCTCGAGACCTTCTCGACTCCCTGGTCACACAGCAACAAATC---CCCAAC 678
Db      208  ValAsnAsnValLeuProValPheAspAsnLeuMetGlnGlnLysLeuValAspGlnAsn 227
QY      679  GTTTTTCACAGCAGATGTGTGGAGCGCGCTGCGCTGCGATCCTGGGACCAACGGA 738
Db      228  IlePheSerPheTyr-----LeuSerArgAspProAspAlaGlnProGly 242
QY      739  GCTACTCTCTCTCTGCTGCTGGAATGACCAAGTTTGTATAAAGGACACATCTGCTATATC 798
Db      243  GlyGluLeuMetLeuGlyThrAspSerLysTyrIleLysGlySerLeuSerTyrLeu 262
QY      799  CCTATTAAAGCAAGCTGGTACTACCAAGATAGAAATTCGAAATGGAAATGGAGCCAA 858
Db      263  AsnValThrArgLysAlaTyrTrpGlnValHisLeuAspGlnValGlnVal---AAsor 281
QY      859  AGCCTTAATCTGGACGTGACAGATATACGCGACAGACAGCCCATCGTGACAGTGCAC 918
Db      282  GlyLeuThrLeu---CysLysGlu-----GlyCysGluAlaIleValAspThrGlyThr 298
QY      919  ACCTCTGCTGCGCTGCGCCAGAGAGTGTGTGATCGGTGGTGGAGAGCTGNGCCCGCA 978
Db      299  SerLeuMetValGlyPro-----ValAspGluValArgGlnLeuGlnLysAlaIle 315
QY      979  TCTCTGATTCACGAATCTCTGATGTTTCTGACTGGGTCCAGCTGGCGTCTGGACG 1038
Db      316  GlyAlaValProLeuIleGlnGlyGluTyr-----MetIleProCys----- 329
QY      1039  AATTCGGAACACCTGGCTTACTCCCTTAAATCTCCATCTACCGAGAGAGAGAAC 1098
Db      330  -----GluLysValSerThrLeuProAlaIleThrLeuLysLeu----- 342
QY      1099  TCCAGCAGGTCATTCGCTATCAACATCTCCCTCAGCTTACATTCAGCCCATGATGG 1158
Db      343  GlyGlyLysGlyTyrLysLeuSer-----ProGluAspTyrThrLeuLysValSerGln 360
QY      1159  GCGCGGCTGAATATGATGATTTACCGATTC-----GCCATTTCCCATCCACAAA 1209
Db      361  AlaGlyLysThrLeuCysLeuSerGlyPheMetGlyMetAspIleProProProSerGly 380
QY      1210  GCCTG---CTGATCGGTGCCAGGTGATGGAGGCTTCTACGTCATCTTCGACAGAGCC 1266
Db      381  ProLeuTyrIleLeuGlyAspValPheIleGlyArgTyrTyrThrValPheAspArgAsp 400
QY      1267  CAGAAAGAGGTGGCTTCGCA 1287
Db      401  AsnAsnArgValGlyPheAla 407

RESULT 15
CATD_MOUSE  STANDARD;  PRT:  410 AA.
AC  P18242;
DT  01-NOV-1990 (Rel. 16, Created)
DT  01-NOV-1990 (Rel. 16, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Cathepsin D precursor (EC 3.4.23.5).
CN  CTSD.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A.

```

```

RC  STRAIN=C57BL/6J; ISSUE=Brain;
RX  MEDLINE=91088345; PubMed=2263503;
RA  Diedrich J.F., Staskus K.A., Retzel E.F., Haase A.T.;
RT  "Nucleotide sequencing of a cDNA encoding mouse cathepsin D.";
RL  Nucleic Acids Res. 18:7184-7184(1990).
RN  [2]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=90326544; PubMed=2374732;
RA  Grusby M.J., Mitchell S.C., Glimcher L.H.;
RT  "Molecular cloning of mouse cathepsin D.";
RL  Nucleic Acids Res. 18:4008-4008(1990).
RN  [3]
RP  SEQUENCE FROM N.A.
RC  STRAIN=C57BL/6J;
RX  MEDLINE=94280822; PubMed=8011168;
RA  Iketani M., Perschl A., Saftig P., von Figura K., Peters C.;
RT  "Mouse cathepsin D gene: molecular organization, characterization of
the promoter, and chromosomal localization.";
RL  DNA Cell Biol. 13:419-427(1994).
CC  FUNCTION: Acid protease active in intracellular protein breakdown.
CC  CATALYTIC ACTIVITY: Specificity similar to, but narrower than,
that of pepsin A. Does not cleave the 4-Gln-His-5 bond in B
chain of insulin.
CC  SUBUNIT: CONSISTS OF A LIGHT CHAIN AND A HEAVY CHAIN.
CC  SUBCELLULAR LOCATION: Lysosomal.
CC  SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
EMBL: X53337; CAA37423.1;
EMBL: X52886; CAA37067.1;
EMBL: X68378; CAA48453.1;
EMBL: X68379; CAA48453.1; JOINED.
EMBL: X68380; CAA48453.1; JOINED.
EMBL: X68381; CAA48453.1; JOINED.
EMBL: X68382; CAA48453.1; JOINED.
EMBL: X68383; CAA48453.1; JOINED.
PIR: I48278; KHM5D.
HSSP: P07339; LLYB.
MEROPS: A01.009; -.
MGD: MGI:88562; Ctcd.
InterPro: IPR001969; Aspprotease_site.
InterPro: IPR001461; AspproteaseA1.
Pfam: PF00026; asp_1.
PRINTS: PR00792; PEPsin.
PROSITE: PS00141; ASP_PROTEASE; 2.
KW  Hydrolase; Aspartyl protease; Glycoprotein; Lysosome; Zymogen; Signal.
FT  SIGNAL 1 20 POTENTIAL.
FT  PROPEP 21 64 ACTIVATION PEPTIDE (POTENTIAL).
FT  CHAIN 65 410 CATHEPSIN D.
FT  ACT_SITE 97 97 BY SIMILARITY.
FT  ACT_SITE 293 293 BY SIMILARITY.
FT  DISULFID 91 16C BY SIMILARITY.
FT  DISULFID 110 117 BY SIMILARITY.
FT  DISULFID 284 288 BY SIMILARITY.
FT  DISULFID 327 364 BY SIMILARITY.
FT  CARBOHYD 134 134 N-LINKED (GLCNAC...) (BY SIMILARITY).
FT  CARBOHYD 261 261 N-LINKED (GLCNAC...) (BY SIMILARITY).
SQ  SEQUENCE 410 AA; 44954 MW; DC4928EC46928BF0 CRC64;

Alignment Scores:
Seq. No.: 5,64e-13 Length: 410
Score: 306.50 Matches: 103
Percent Similarity: 44.53% Conservative: 64
Best Local Similarity: 27.47% Mismatches: 123
Query Match: 9.18% Indels: 85
DB: 1 Gaps: 15

```


US-09-806-194-1 (1-1804) x CATD_MOUSE (1-410)

```
QY 274 TACTACGTGAGATGCTGCGGACCCGCCAGAGACTACAGATTCTCGTTGACACT 333
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 79 TyrTyrGlyAspIleGlyIleGlyThrProGlnCysPheThrValPheAspIle 58
QY 334 GGAAGCAGTAACCTTCCGCTGGCAGGACCCGCCACTCTACATACACACAGC 384
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 99 GlySerSerAsnLeuTrpValProSerIleHisCysLysIleLeuAspIleAlaCysTrp 118
QY 385 -----TACTTTACACAGAGGTCTAGCACATACCGCTCCCAAGGCGCTTACGTC 435
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 119 ValHisLysTyrAsnSerAspLysSerSerThrTyrValLysAsnGlyThrSerPhe 138
QY 436 ACAGTGAAGTACACACAGAGCAAGCTGGAGGGCTTCGTTGGGAGACACCTGTCACCATC 495
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 139 AspIleHisTyrGlySerGlySerLeuSerGlyTyrLeuSerGlnAspThrValSerVal 158
QY 496 CCCAAAGGCTCAATACTCTTTCTTCTGTCACCAATTGCCACT-----ATTTTGAA 546
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 159 ProcLysSerAspGlnSerLysAlaArgGlyIleLysValGluLysGlnIlePhe--- 177
QY 547 TCACAGAAATTTCTTTTGCTGGGATT-----AAATCGAATCGAAACTT 591
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 178 GlyGluAlaThrLysGlnProGlyIleValPheValAlaAlaLysPheAspGlyIleLeu 197
QY 592 GGCTAGCTTATGCCACTTGCACAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCC 651
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 198 GlyMetGlyTyrProHisIleSer-----ValAsnAsnValLeuProValPheAspAsn 215
QY 652 CTGGTGACACAAACAACATC-----CCCAACGTTTCTCCATCCAGATGTGTGGAGCCGCC 708
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 216 LeuMetGlnGlnLysLeuValAspLysAsnIlePheSerPheTyr----- 230
QY 709 TTGCCCGTTCTCGATCTGGACCAAGGAGGTAGTCTTCTCTGGTGGAAATCGAACCA 768
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 231 LeuAsnArgAspProGluGlyGlnProGlyGlyGluLeuMetLeuGlyGlyThrAspSer 250
QY 769 AGTTTGATTAAGGACATCTGGTATACCCCTATTAAAGCAAGAGTGTTACTACCAGATA 828
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 251 LysTyrTyrHisGlyGluLeuSerTyrLeuAsnValThrArgLysAlaTyrTrpGlnVal 270
QY 829 GAAATCTGAATTTGAATTTGGAGCCCAAGCCTTAATCTGACTGACAGAGATATAAC 888
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 271 HisMetAspGlnLeuGluValGlyAsnGlu---LeuThrLeu---CysLys-----Gly 286
QY 889 GCAGACAAGGCCCTGCTGGACAGTGGCACACAGCTGCTGGGCGCTGCCCCAGAGCTGTT 948
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 287 GlyCysGluAlaIleValAspThrGlyThrSerLeuLeuValGlyProValGluGluVal 306
QY 949 GATCGCGTGTGGAACTGTGGCCCGCGCATCTCTGATTCAGAAATCTCTGAIGGTTTC 1008
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 307 LysGluLeuGlnLysAlaIleGlyAlaValProLeuIle----- 319
QY 1009 TGCAGTGGGTCCCGAGTGGCGCTCTCGAGCAATTCGGAACACCTTGGTCTTACTTCCT 1068
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 319 ----- 319
QY 1069 AAAATCTCCATCTACCTGAGAGATGAGAACTCCAGCAGGTCATTCGATACAAATCTGT 1126
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 320 -----GlnGlyGluTyrMetIleProCysGluLysValSerSerLeu 333
QY 1129 CCTCAGCTTTACATTCAGGCCATGATGGGGCGCGCTGAATATGATGT----- 1179
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 334 ProThrValTyrLeuLys-----LeuGly---GlyLysAsnTyrGluLeuHisProAsp 350
QY 1180 ---TACCGATTCCGCATTCC----- 1197
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 351 LysTyrIleLeuLysValSerGlnGlyLysThrIleCysLeuSerGlyPheMetGly 370
QY 1198 -----CCATCCCAAAATGCGCTGGTGCATCGGTGCCACCGGTGATGGAGGC 1242
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
DB 371 MetAspIleProProProSerGlyProLeuTrpIleLeuGlyAspValPheIleGlySer 390
QY 1243 TTCTACGTCTATCTTTCGACAGAGCCCGAGAGAGGGTGGGCTTCGCA 1287
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 391 TyrTyrThrValPheAspArgAspAsnArgValGlyPheAla 405
```

Search completed: September 25, 2003, 11:01:09
Job time : 56.5 secs

Result No.	Score	Query Match	Length	DB ID	Description
1	1178.5	35.3	501	A509090	aspartic protease
2	367.5	11.0	383	JC75753	pepsinogen (EC 3.4.21.1)
3	363.5	10.9	377	PEMOCJ1	gastricsin
4	385.5	10.6	384	A39334	gastricsin
5	355	10.6	389	JE0371	pepsin C (EC 3.4.21.1)
6	353	10.6	388	A29937	gastricsin
7	351.5	10.5	382	JC7246	pepsinogen (EC 3.4.21.1)
8	324.5	9.7	394	B43356	gastricsin
9	320	9.6	385	JC75755	pepsinogen (EC 3.4.21.1)
10	320	9.6	402	1 REMSK	renin (EC 3.4.21.1)
11	313.5	9.4	509	2 S66536	oryzasin (EC 3.4.21.1)
12	313	9.4	392	1 A24608	gastricsin
13	310	9.3	383	2 A41443	pepsin (EC 3.4.21.1)
14	308.5	9.2	412	1 KHHUD	cathepsin D

DB:

Alignment Scores:			
Pred. No.:	1.28e-76	Length:	501
Score:	1178.50	Matches:	240
Percent Similarity:	62.04%	Conservative:	82
Best Local Similarity:	46.24%	Mismatches:	164
Query Match:	35.22%	Indels:	33
DB:	2	Gaps:	9

QY	1069	AAAATCTCCATCTACCTGACGAGATGAGAACTCCAGCAGGTCATCTCCGTATCAACAATCCTG	1120
Db		:	
QY	1129	CCTCAGCTTTACATTTACGCCCATGATGGGGCGGCGCTGAAATAT---	1185
Db		: :	
QY	1245	TTGGCGCATTTCCCGCATCCCAAAATGCCCTGGTGATCGGTGCCAGCGGTGATGGAGGGCTTC	1300
Db		:	
QY	1305	TAGCTCATCTTCGACAGAGCCCAAGACAGAGGTGGCTTCGACAGGACCCCTGTGCGAAG	1360
Db		: : :	
QY	1422	TyrValValPheAspArgAlaArgLysArgIleGlyPheAlaValSerHisVal	1486
Db		: :	
QY	1486	TGTCAGGTGCTGCAGTGCTGAAATTTCCGGGCGCTTCTCAACAGAGGATGTAGCCAGC	1550
Db		: :	
QY	1550	AACTGTGTCGCCGCTCAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCTATGCGCTC	1614
Db		:	
QY	1614	ATGAGCGTCTGTGAGCGCATCTCCTCTTCTTAAATCGTCTGCTGCTGCGCCGCTTCGG	1678
Db		:	
QY	1678	AlaAlaIleCys---AlaLeuPheMetLeuProLeuMetValCysGlnTrpArg	1742
Db		:	
QY	1742	TGTCAGGTGCTGCAGTGCTGAAATTTCCGGGCGCTTCTCAACAGAGGATGTAGCCAGC	1806
Db		: :	
QY	1806	AACTGTGTCGCCGCTCAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCTATGCGCTC	1870
Db		:	
QY	1870	ATGAGCGTCTGTGAGCGCATCTCCTCTTCTTAAATCGTCTGCTGCTGCGCCGCTTCGG	1934
Db		:	
QY	1934	AlaAlaIleCys---AlaLeuPheMetLeuProLeuMetValCysGlnTrpArg	2000
Db		:	
QY	2000	TGTCAGGTGCTGCAGTGCTGAAATTTCCGGGCGCTTCTCAACAGAGGATGTAGCCAGC	2064
Db		: :	
QY	2064	AACTGTGTCGCCGCTCAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCTATGCGCTC	2128
Db		:	
QY	2128	ATGAGCGTCTGTGAGCGCATCTCCTCTTCTTAAATCGTCTGCTGCTGCGCCGCTTCGG	2192
Db		:	
QY	2192	AlaAlaIleCys---AlaLeuPheMetLeuProLeuMetValCysGlnTrpArg	2258
Db		:	
QY	2258	TGTCAGGTGCTGCAGTGCTGAAATTTCCGGGCGCTTCTCAACAGAGGATGTAGCCAGC	2322
Db		: :	
QY	2322	AACTGTGTCGCCGCTCAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCTATGCGCTC	2386
Db		:	
QY	2386	ATGAGCGTCTGTGAGCGCATCTCCTCTTCTTAAATCGTCTGCTGCTGCGCCGCTTCGG	2450
Db		:	
QY	2450	AlaAlaIleCys---AlaLeuPheMetLeuProLeuMetValCysGlnTrpArg	2516
Db		:	
QY	2516	TGTCAGGTGCTGCAGTGCTGAAATTTCCGGGCGCTTCTCAACAGAGGATGTAGCCAGC	2580
Db		: :	
QY	2580	AACTGTGTCGCCGCTCAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCTATGCGCTC	2644
Db		:	
QY	2644	ATGAGCGTCTGTGAGCGCATCTCCTCTTCTTAAATCGTCTGCTGCTGCGCCGCTTCGG	2708
Db		:	
QY	2708	AlaAlaIleCys---AlaLeuPheMetLeuProLeuMetValCysGlnTrpArg	2774
Db		:	
QY	2774	TGTCAGGTGCTGCAGTGCTGAAATTTCCGGGCGCTTCTCAACAGAGGATGTAGCCAGC	2838
Db		: :	
QY	2838	AACTGTGTCGCCGCTCAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCTATGCGCTC	2902
Db		:	
QY	2902	ATGAGCGTCTGTGAGCGCATCTCCTCTTCTTAAATCGTCTGCTGCTGCGCCGCTTCGG	2966
Db		:	
QY	2966	AlaAlaIleCys---AlaLeuPheMetLeuProLeuMetValCysGlnTrpArg	3032
Db		:	
QY	3032	TGTCAGGTGCTGCAGTGCTGAAATTTCCGGGCGCTTCTCAACAGAGGATGTAGCCAGC	3096
Db		: :	
QY	3096	AACTGTGTCGCCGCTCAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCTATGCGCTC	3160
Db		:	
QY	3160	ATGAGCGTCTGTGAGCGCATCTCCTCTTCTTAAATCGTCTGCTGCTGCGCCGCTTCGG	3224
Db		:	
QY	3224	AlaAlaIleCys---AlaLeuPheMetLeuProLeuMetValCysGlnTrpArg	3290
Db		:	
QY	3290	TGTCAGGTGCTGCAGTGCTGAAATTTCCGGGCGCTTCTCAACAGAGGATGTAGCCAGC	3354
Db		: :	
QY	3354	AACTGTGTCGCCGCTCAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCTATGCGCTC	3418
Db		:	
QY	3418	ATGAGCGTCTGTGAGCGCATCTCCTCTTCTTAAATCGTCTGCTGCTGCGCCGCTTCGG	3482
Db		:	
QY	3482	AlaAlaIleCys---AlaLeuPheMetLeuProLeuMetValCysGlnTrpArg	3548
Db		:	
QY	3548	TGTCAGGTGCTGCAGTGCTGAAATTTCCGGGCGCTTCTCAACAGAGGATGTAGCCAGC	3612
Db		: :	
QY	3612	AACTGTGTCGCCGCTCAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCTATGCGCTC	3676
Db		:	
QY	3676	ATGAGCGTCTGTGAGCGCATCTCCTCTTCTTAAATCGTCTGCTGCTGCGCCGCTTCGG	3740
Db		:	
QY	3740	AlaAlaIleCys---AlaLeuPheMetLeuProLeuMetValCysGlnTrpArg	3806
Db		:	
QY	3806	TGTCAGGTGCTGCAGTGCTGAAATTTCCGGGCGCTTCTCAACAGAGGATGT	

```
Db 20 -----ValProLeuLysLysPheLysSerMetArgGluValMetArgGluAsp 35
QY 121 CGGGTGGTGGCCCGCCGAGCCGCGGAGCCCTGCGGAGCGCCACCGCCAGCGCTTG 190
Db 36 GlyLysAlaProLeuValAspProAlaThr-----LysTyrTyrAsnGlnTyr 52
QY 181 CGCGTCCGCTGGAGCTGCGCTGGCGTCCCGCGGCGCGCCCAACTCTTGGCCATG 240
Db 53 AlaThrAlaTyrGluPro-----LeuSerAsnTyrMetAspMet 65
QY 241 GTACAGAACTCGAGGGGACTCTGCGCGGCTACTACTCGAGATGCTGATCGGAGCC 300
Db 66 -----SerTyrGlyGlySerIleGlyThr 75
QY 301 CCGCCGAGAGCTACAGATTCCTGGTGGAGGAGGAGTAACCTTGGCGTGGCAGGA 360
Db 76 ProProGlnAsnPheLeuValLeuPheAspThrGlySerSerAsnLeuIrpValAlaSer 95
QY 361 ACCCGGACTCTACATAGACAG-----TACTTTGACACAGAG 399
Db 96 Thr-----TyrCysGlnSerGlnAlaCysThrAsnHisProLeuPheAsnProSer 422
QY 400 AGGTGACACATACCGCTCCAAAGGCTTGGAGTCCACAGTGAAGTACACACAAAGAAC 459
Db 113 GlnSerSerThrTyrSerSerAsnGlnGlnPheSerLeuGlnTyrGlyThrGlySer 132
QY 460 TGGACGGGCTTCGTTGGGAGAGCTCGTCCACATCCCAAGAGGCTTCAATCTCTTT 519
Db 133 LeuThrGlyIleLeuGlyTyrAspThrValIleGln----- 145
QY 520 CTGTGCAACATCCCACTATTTTGAATCAGAGAAATCTPTTTTG-----CCG 567
Db 146 -----AsnValAla-----IleSerGlnGlnGluPheGlyLeuSerGluThrGluPro 161
QY 568 GGG-----ATTAAATGAATGAATACITTCGCTAGCTATGATGCACACTT 612
Db 162 GlyThrAsnPheValTyrAlaGlnPheAspGlyIleLeuGlyLeuAlaTyrProSerIle 181
QY 613 GCCAAGCCATCAAGTCTCTGGAGACCTTCTCGACTCCCGTGGTACACAAAGCAATC 672
Db 182 AlaValGlyGlyAlaThr-----ThrValMetGlnGlyMetMet-----GlnGlnAsnLeu 198
QY 673 CCGAAC-----GTTTCTCCATCCAGATGTGGAGCGCGCTGCGCTGCGATCT 726
Db 199 LeuAsnGlnProIlePheGlyPheTyrLeuSerGlyGln-----Ser 212
QY 727 GGGACCAACGGAGTAGTCTGTCTTGGTGGAAATTGAACCAAGTTGTGATAAGAGAG 786
Db 213 SerGlnAsnGlyGlyGluValAlaPheGlyValAspAlaAsnTyrTyrThrGlyGln 232
QY 787 ATCTGGTATACCCCTATTAGAGAGAGTGTACTACAGATAGAAATCTGGAATTCGAA 846
Db 233 IleTyrTrpThrProValThrSerGluThrTyrTrpGlnIleGlyIleGlnGlyPheSer 252
QY 847 ATTGGAGGCAAGCCCTTAATCTGGACTCGAGAGTATAACACAGACAGGCGCATCG 906
Db 253 IleAsnGlyGlnAlaThrGlyTyr-----CysSerGln-----GlyCysGlnAlaLeuVal 269
QY 907 GACAGTGGCAGCGCTGCTGCGCGCTGCCCGCCAGAGGCTGTTGATCGGTGCGAGCT 966
Db 270 AspThrGlyThrSerLeuLeuThrAlaProGlnSerValPheSerSerLeuIleGlnSer 289
QY 967 GTGCGCCGCGCATCTCTGATCCAGAAATCTCTGATGGTTCCTGGACTGGTCCCACTG 1026
Db 290 IleGly-----AlaGlnGlnAspGlnAsnGlnTyrVal 301
QY 1027 CGGTGCTGGAGAAATCGGAACACCTTGCTCTACTTCCCTAAATCTCCTCACTACCTG 1086
Db 302 ValSerCysSerAsnIleGlnAsn-----LeuProThrIleSerPheThrIle 317
QY 1087 AGAGATGAGAACTCCAGAGGCTATTCGATATCACAATCTCGCTCAGCTTTTACATTT 1143
```

```
Db 313 -----SerGlyValSerPheProLeuPro-----ProSerAlaTyrValLeu 331
QY 1144 CAGCCCATGATGGCGCGCGCTGAATATGATGATGATGATGATGATGATGATGATGAT 1197
Db 332 GlnGlnSerSerGly-----TyrCys-----ThrIleGlyIleMetProThr 345
QY 1198 -----CATCCACAAATGCG-----CTGGTATCGGTGCCAGGTGATGCGAGGC 1242
Db 346 TyrLeuProSerGlnAsnGlyGlnProLeuIrpIleLeuGlyAspValPheLeuArgGlu 365
QY 1243 TTCTACGTCATCTTCGACAGAGAGGAGGTGGCTTCGAGAGGAGC 1293
Db 366 TyrTyrSerValTyrAspLeuGlyAsnGlnValGlyPheAlaThrAla 382
RESULT 3
PEMQCJ
gastricsin (EC 3.4.23.3) precursor - Japanese macaque (fragment)
N.Alternate names: pepsin C
C.Species: Macaca fuscata (Japanese macaque)
C.Date: 13-Aug-1986 #sequence_revision 19-Oct-1995 #text_change 18-Jun-1999
C.Accession: S19683; A00986; A22402; S16066
R.Kageyama, T.; Tanabe, K.; Koiwai, O.
Eur. J. Biochem. 202, 205-215, 1991
A.Title: Development-dependent expression of isozymogens of monkey pepsinogens and st
A.Reference number: S19681; MUID:92037645; PMID:1935977
A.Accession: S19683
A.Molecule type: mRNA
A.Residues: 1-377 <KAG>
A.Cross-references: EMBL:X59754; NID:g38072; PIDN:CAA42426.1; PID:g38073
R.Kageyama, T.; Takahashi, K.
J. Biol. Chem. 261, 4406-4419, 1986
A.Title: The complete amino acid sequence of monkey progastricsin.
A.Reference number: A00986; MUID:86168133; PMID:3514597
A.Accession: A00986
A.Molecule type: protein
A.Residues: 6-330,'V',332-349,'VY',350-377 <KA2>
R.Kageyama, T.; Takahashi, K.
J. Biochem. 97, 1235-1246, 1985
A.Title: Monkey pepsinogens and pepsins. VII. Analysis of the activation process and
pepsinogens.
A.Reference number: A22402; MUID:85289106; PMID:3928607
A.Accession: A22402
A.Molecule type: protein
A.Residues: 6-65 <KA3>
C.Comment: This enzyme has more restricted specificity than pepsin A.
C.Comment: The enzyme is activated in a two-step process that gives rise to two end p
C.Superfamily: pepsin
C.Keywords: aspartic proteinase; gastric juice; hydrolase; protein digestion; stomach
F.1-5/Domain: signal sequence (fragment) #status predicted <SIG>
F.6-377/Product: progastricsin #status experimental <2YM>
F.6-45/Domain: activation peptide #status experimental <APT>
F.46-377/Product: Gly-gastricsin #status experimental <MIN>
F.49-377/Product: Ser-gastricsin #status experimental <MAT>
F.31-32/Cleavage site: Phe-Leu (pepsin) #status experimental
F.45-46/Cleavage site: Leu-Gly (pepsin) #status experimental
F.48-49/Cleavage site: Leu-Ser (pepsin) #status experimental
F.80.265/Active site: Asp #status predicted
F.93-98,256-260,299-332/Disulfide bonds: #status experimental
```

Alignment Scores:

Pred. No.:	3.87e-18	Length:	377
Score:	363.50	Matches:	118
Percent Similarity:	44.85%	Conservative:	65
Best Local Similarity:	28.92%	Mismatches:	118
Query Match:	10.89%	Indels:	107
DB:	1	Gaps:	19

US-09-806-194-1 (1-1804) x PEMQCJ (1-377)

```
QY 166 CAGCGCGAGCGGTTCGGCGCTGCCCTGAGAGCTGCCCTGGCTGCCCGCGCGCGGCC 225
Db 44 HisPheGlyAspLeuSerValSerTyrGluPro----- 54
```

```

QY 226 AACTTCTGGCCATGGTAGACAACTGCAGGGGACTCTGGCCGCGCTACTACCTGGAG 285
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
55 -----MetAlaTyrMetAsp-----AlaAlaTyrPheGlyGlu 65
QY 286 ATGCTGATCGGACCCCGCAGAGACTACAGATTCTGTTGCACACTGGAAGCAGTAAC 345
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
56 IleSerIleGlyThrProProGlnAsnPheLeuValLeuPheAspThrGlySerSerAsn 85
QY 346 TTGCCCCGTG-----GCAGGAACCCCGCACTCCATACATAGAC 381
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
86 LeuTrpValProSerValTyrCysGlnSerGlnAlaCysThrSerHisSer----- 102
QY 382 ACGTACTTGACACAGAGGCTTAGCACATACCGCTCCAAGGCTTGCAGCTCACAGNG 441
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
103 ---ArgPheAsnProSerGluSerThrTyrSerThrAsnGlnGlnThrPheSerIleu 121
QY 442 AAGTACACAAAGAGCTGAGGGCTGTTGGGGAAGACTCGTCCACCATCCCAAA 501
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
122 GlnTyrGlySerGlySerLeuThrGlyPhePheGlyTyrAspThrLeuThrVal----- 139
QY 502 GGCTTCATACTTCTTCTGTCAACATTGCCACTATTTTGAATCAGAGAAITCTTT 561
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
140 -----GlnSerIleGlnValProAsnGlnGluPheGlyLeuSerGluAsn----- 154
QY 562 TTGCCCTGG-----ATTAAATGGAATGGAATACTTGGCTAGCTTATGCC 606
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
155 GluProGlyThrAsnPheValTyrAlaGlnPheAspGlyIleMetGlyLeuAlaTyrPro 174
QY 607 ACATTGCCAAGCCAYCAAGTTCTCTGGAGACTTCTTGCAGTCCCTGGTGACACAGCA 666
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
175 ThrLeuSerValAspGlyAlaThr-----ThrAlaMetGlnGlyMetValGlnGly 192
QY 667 AACATCCCCAAC---GTTTCTCCATGCAGATGTGGAGCCGCTTGCCGCTGTGTGCA 723
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
193 AlaLeuThrSerProIlePheSerValTyrLeuSerAspGln----- 206
QY 724 TCTGGGACCAAGGAGTAGTCTTGTCTGGTGGGAATGAACCAAGTTGTATAAAGA 783
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
207 GlnGlySerSerGlyGlyAlaValAlaPheGlyGlyValAspSerSerLeuTyrThrGly 226
QY 784 GACATCTGGTAACCCCTATTAAAGGAAGAGTGTACTACCATAGATAAAATCTGAAATG 843
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
227 GlnIleTyrTrpAlaProValThrGlnGluLeuTyrTrpGlnIleGlyIleGluGluPhe 245
QY 844 GAAATTTGGAGCCCAAGCCCTTAATCTGGACTGCAGAGAGTATACCCAGCAAGGCCATC 903
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
247 LeuIleGlyGlyGlnAlaSerGlyTyr---CysSerGlu-----GlyCysGlnAlaIle 263
QY 904 GTGCAGAGTGGCACCAGCTGCTGCGCTGCCAGAGGTTTGTATGGTGGTGGAA 963
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
264 ValAspThrGlyThrSerLeuLeuThrValProGlnGlnTyrMetSerAlaLeuLeuGln 283
QY 964 CTGTGTGCCCGCGCATCTCTGATTCCAGAAATCTCTGATGGTTTCTGGACTGGGCCCCAG 1023
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
284 Ala-----ThrGlyAlaGln 288
QY 1024 CTGCGTGTGGAGCAATTCGGAAACACCTTGTCTTACTTC----- 1065
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
289 -----GluAspGluTyrGlyGlnPheLeuValAlaAsnCysAsnSer 301
QY 1066 -----CCTAAATCTCCATCTACTCTCAGAGATGAGAACTCCAGCAGGTCAITC 1113
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
302 IleGlnAsnLeuProThrLeuThrPheIleIle-----AsnGlyVal 315
QY 1114 CGTATCAAAATCTCGCTCAGCTTTTACATTACGCCCATGATGGGCGCGCTGAATTAT 1173
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
316 GluPheProLeuProSerSerTyrIle-----LeuAsnAsn 328
QY 1174 GAATGTTAC---CGATTCGGCATTTCCCA-----TCCACAAT 1209
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
329 AsnGlyTyrCysThrValGlyValGluProThrTyrLeuSerAlaGlnAsnSerGlnPro 348
QY 1210 CGCGTGTGTGATCGTGCCAGGCTGATGGAGGGCTTCTACGTCTCTTCGACAGAGCCCA 1269

```

```

Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
349 LeuTrpIleLeuGlyAspValPheLeuArgSerTyrTyrSerValTyrAspLeuSerAsn 368
QY 1270 AAGAGGTGGGCTTCGCGAGCGAGC 1293
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
369 AsnArgValGlyPheAlaThrAla 376
RPSULT 4
A39314
#StraticsIn (EC 3.4.23.3) precursor - bullfrog
C:Species: Rana catesbeiana (bullfrog)
C:Date: 19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change 22-Jun-1999
C:Accession: A39314
R:Yakabe, E.; Tanji, M.; Ichinose, M.; Goto, S.; Miki, K.; Kurokawa, K.; Ito, H.; Kag
J. Biol. Chem. 266, 22436-22443, 1991
A:Title: Purification, characterization, and amino acid sequences of pepsinogens and
A:Reference number: A39314; MUID:92042186; PMID:1939266
A:Accession: A39314
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-384 <YAK>
A:Cross-references: GB:M73750; NID:g213687; PIDN:AAA49530.1; PID:g213688
C:Superfamily: pepsin
C:Keywords: aspartic proteinase; hydrolase; protein digestion

Alignment Scores:
Pred. No.:      Length:      384
Score:          355.50      Matches:      120
Percent Similarity: 42.70%      Conservative: 73
Best Local Similarity: 26.55%      Mismatches:   136
Query Match:      10.65%      Indels:       123
DB:              2          Gaps:         21

US-09-806-194-1 (1-1804) x A39314 (1-384)
QY 67  GAGTGGCCCGCGCGCCTTACGCTGCCCTCGGCTCGGCTGGCGCGGCCAGAACCGGTA 126
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
12  GlnLeuSerGluGlyIleIleLysValProLeuLysLysPheLysSerMetArgGluVal 31
QY 127 GTT-----GCGCCACCCCGGAGCCCGGACCCCTCGCGAGCCGAC 168
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
32 MetArgAspHisGlyIleLysAlaProValValAspProAlaThr-----LysTyr 48
QY 169 GCGAGCGCTTGGCGCTCGCCCTCGAGCCTGCCCTGCCCTGCCCGCGCGCGCGCAAC 228
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
49  TyrAsnAsnPheAlaThrAlaPheGluPro-----LeuAlaAsn 61
QY 229 TTCCTGGCCATGTAGACAACTCCAGGGGACTCTCGCGCGGCTACTACTCTGGAGATG 288
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
62  TyrMetAspMet-----SerTyrTyrGlyGluIle 71
QY 289 CTGATCGGGAGCCCGCGCAGAGCTACAGATTCTCTTGACACTGGAAGCAGTAACITT 348
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
72  SerIleGlyThrProGlnAsnPheLeuValLeuPheAspThrGlySerSerAsnLeu 91
QY 349 GCGTGGCAGAACCCCGCACTCTCTACATAGACAG-----TAC 387
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
92  TrpVal-----ProSerThrTyrCysGlnSerGlnAlaCysThrAsnHisProGln 108
QY 388 TTTGACACAGAGAGGTCTAGCACATACCGCTCCCAAGGCTTTCAGCTGCACAGTGAAGTAC 447
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
109 PheAsnProSerGlnSerSerSerTyrSerSerAsnGlnGlnGlnPheSerLeuGlnTyr 128
QY 448 ACACAAGAAAGCTGGAGGGCTCGTTGGGGAAGACCTCGTCCACCATCCCAAGGCTTC 507
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
129 GlyThrGlySerLeuThrGlyIleLeuGlyTyrAspThrValGlnIleGln----- 145
QY 508 AATACTCTCTTCTGTGCACATGGCCACTATTTTGAATCAGAGAAATTTCTTTTG--- 564
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
146 -----AsnIleAla-----IleSerGlnGlnGluPheGlyLeuSer 157
QY 565 -----CCTGG-----ATTAATGGAATGGAATACTTGGCTAGCT 600
Db      :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||

```

```
Db 158 ValThrGluProGlyThrAsnPheValTyrAlaGlnPheAspGlyLeuGlyLeuAla 177
QY 601 TAIGCCACACTGCCAAGCCATCAAGTTCCTGGAGACCTCTTCGACACCTCGTGGACA 660
Db 178 TyrProSerIleAlaGluGlyGcYAlaThr-----ThrValMetGlnGlyMetIle--- 194
QY 661 CAAGCAACATCCCAAC-----GTTTCTCCATGCAGATGTGTGGACCGGCTTCGAC 714
Db 195 GlnGlnAsnLeuIleAsnGlnProLeuPheAlaPheTyrLeuSerGly----- 210
QY 715 GTTGCTGGATCTGGGACCAACGAGGTAGTCTTGTCTGGTGGGAATTGAACCAAGTTG 774
Db 211 ---GlnGlnAsnSerGlnAsnGlyGlyGluValAlaPheGlyGlyValAspGlnAsnTyr 229
QY 775 TATAAAGGACACTGGTATACCCCTATTAAAGAAAGAGTGGTACTACCATAGAAATT 834
Db 230 TyrSerGlyGlnIleTyrTrpThrProValThrSerGluThrTyrTrpGlnIleGlyIle 249
QY 835 CTGAATTTGGAATTTGGAGCCCAAGCCTTAATCTGGACTGGACAGATATACGCGAGAC 894
Db 250 GlnGlyPheSerValAsnGlnGlnAlaThrGlyTyr---CysSerGln-----GlyCys 266
QY 895 AAGGCCATCTGTGACAGTGGCACACGCTGCTCGGCTGCCCCAGAAAGTGTGTGATCG 954
Db 267 GlnGlyIleValAspThrGlyThrSerLeuLeuThrAlaProGlnSerValPheSerSer 286
QY 955 GTGGTGAAGCTGTGGCCCGCGCATCTCTGATTCAGAAATCTCTGATGTTTCTGGACT 1014
Db 287 LeuMetGlnSerIle----- 291
QY 1015 GGTGCCAG-----CTGGCGTGTGGAGAAATTCGGAACA 1050
Db 292 GlyAlaGlnGlnAspGlnAsnGlyGlnTyrAlaValSerCysSerAsnIleGlnSerLeu 311
QY 1051 CCTTGGTCTTACTCCCT-----AAAATCTCCATCTACTCT 1086
Db 312 ProThrIleSerPheThrIleSerGlyValSerPheProLeuProSerAlaTyrVal 331
QY 1087 AGAGATAGAACTCC-----AGCAGGTGATTCGGTATCACAATCCCTGGCTACGCTT 1137
Db 332 LeuGlnGlnAsnSerGlyTyrCysThrIleGlyIleMetProThrTyrLeuProSerGln 351
QY 1138 TACATTACGCCCATGTAGGGGGCGGCGCTCAATTATGAATGTACCATTCGGCATTC 1197
Db 352 AsnGlyGlnProLeuTrp----- 357
QY 1198 CCATCCACAAATCGCTGGTGATCGGTGCACCGGTGATGGAGGCTTCTACGTATCTTC 1257
Db 358 -----IleLeuGlyAspValPheLeuArgGlnTyrTyrSerValTyr 371
QY 1258 GACAGAGCCCAAGAGAGGTGGCTTCGACGCGAGC 1293
Db 372 AspLeuGlyAsnAsnGlnValGlyPheAlaAlaAla 363

RESULT 5
JE0371
pepsin C (EC 3.4.23.-) precursor - chicken
N:Alternate names: pepsinogen C
C:Species: Gallus gallus (chicken)
C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 11-May-2000
C:Accession: JE0371
R:Sakamoto, N.; Saiga, H.; Yasugi, S.
Biochem Biophys Res Commun. 250, 420-424, 1998
A:Title: Analysis of temporal expression pattern and cis-regulatory sequences of chicken
A:Reference number: JE0370; MUID:98440813; PMID:9753645
A:Accession: JE0371
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-389 <SAK>
C:Superfamily: pepsin
C:Keywords: aspartic proteinase; hydrolase

Alignment Scores:
```

```
Pred. No.: 1:58e-17 Length: 389
Score: 355.00 Matches: 114
Percent Similarity: 43.32% Conservative: 58
Best local Similarity: 28.72% Mismatches: 121
Query Match: 10.63% Indels: 104
DB: 2 Gaps: 16

US-09-806-194-1 (1-1804) x JE0371 (1-389)

QY 223 GCCAACTTCTGGCCATGGTAGAACCTCGAGGGGACTCTGGCCGGGCTACTACCTG 282
Db 56 SerAsnPheAlaThrAlaTyrGluProLeuAlaAsnAsnMetAspMetSerTyrTyrGly 75
QY 283 GAGATGCTGATCGGACCCGCCCGCAGAGTACAGATTCTCTGTGACATCGGAAGCAGT 342
Db 76 GluIleSerIleGlyIleThrProGlnAsnPheLeuValLeuPheAspThrGlySerSer 95
QY 343 RACTTTGCGGTGGCAGGAACC-----CCGACCTCCTACATA 378
Db 96 AsnLeuTrpValProSerThrLeuCysGlnSerGlnAlaCysAlaAsnHisAsn----- 113
QY 379 GACAGCTACTTTGACACAGAGAGGTCTAGCACATACCGCTCCCAAGGGCTTTGACGTGACA 438
Db 114 -----GluPheAspProAsnGluSerThrPheSerThrGlnAspGluPhePheSer 131
QY 439 GTGAAGTACACAAAGAGAGCTGGACGGCTTCGTGGGAAGACCTCGTCACCAATCCCC 498
Db 132 LeuGlnTyrGlySerGlySerLeuThrGlyIlePheGlyPheAspThrValThrIle--- 150
QY 499 AAAGCGCTCAAT-----ACTCT 516
Db 151 GlnGlyIleSerIleThrAsnGlnGluPheGlyLeuSerGluThrGluProGlyThrSer 170
QY 517 TTCTTTGTCACAAITGCCACTATTTTGAATCAGAGAATTTCTTTTCCCTGGGATTAAA 576
Db 171 PheLeuTyrSer-----Pro 175
QY 577 TGGATGGAACTTGGCTAGCTTATGCCACACTTGCCAAAGCCATCAAGTCTCTGGAG 636
Db 176 PheAspGlyIleLeuGlyLeuAlaPheProSerIle-----SerAlaGlyGlyAla 192
QY 637 ACCTTCTTCGACTCCCTGTGACACAAGCAACATCCCAAC-----GTTTCTCCCAIG 690
Db 193 ThrThrValMetGlnLysMetLeuGlnGluAsnLeuLeuAspPheProValPheSerPhe 212
QY 691 CAGATGTGTGGACCGGCTTGCCTGTCGTGGACCAACGAGGAGTAGTCTTGTCT 750
Db 213 TyrLeuSerGlyGln-----GluGlySerGlnGlyGlyGluLeuVal 226
QY 751 TTGGTGGAAATTTGAACCAAGTTTGTATAAGGAGACATCTGTATACCCCTATTAAAGGAA 810
Db 227 PheGlyGlyValAspProAsnLeuTyrThrGlyGlnIleThrTrpThrProValThrGln 246
QY 811 GAGTGTACTACCATAGAAATCTGAAATTTGGAATTTGGAGCCCAAGGCTTAAATCTG 870
Db 247 ThrThrTyrTrpGlnIleGlyIleGluAspPheAlaValGlyGlyGlnSerSerGlyTrp 266
QY 871 GACTCGACAGAGATTAACGACAGACAGGCCCATCTGTGGACAGTGGCCACACCGCTGCTGCGC 930
Db 267 ---CysSerGln-----GlyCysGlnGlyIleValAspThrGlyThrSerLeuLeuThr 283
QY 931 CTGCCCCAGAGGCTTTTGTATCGGTGTGGAGGTGTGGCCCGGCATCTCTGATTCCA 990
Db 284 ValProAsnGlnValPheThrGluLeuMetGlnTyrIleGly-----Ala 298
QY 991 GAATTCTCTGATGTTTCTTGGACTGGTCCCGCCAGCTGGCGTGGACGAATTCGGAACA 1050
Db 299 GlnAlaAspAsp-----SerGlyGlnTyrValAlaSerCysSerAsnIleGlu--- 314
QY 1051 CCTTGGTCTTACTTCCCTAAATC----- 1074
Db 315 -----TyrMetProThrIleThrPheValIleSerGlyThrSerPheProLeuPro 331
```

```

F:50-388/Product: gastricsin #status experimental <MAT>

Alignment Scores:
Pred. No.:      2.2e-17      Length:      388
Score:          353.00      Matches:    120
Percent Similarity: 44.79%    Conservative: 65
Best Local Similarity: 29.06%  Mismatches: 120
Query Match:     10.57%     Indels:     108
DB:              2          Gaps:         21

US-09-806-194-1 (1-1804) x A29937 (1-388)

QY 154 CCTCGCAGGCCACCGCCGACGGC---TTGGCGCTCGCCTGGAGCCTGCCTGGCGTCC 210
Db 50 ProAlaTrpIysIyrArgPheGlyAspLeuSerValThrTyrGluPro----- 65

QY 211 CCCGGGGCGCGCCCAACTTCTTTGGGCATGGTAGACAACCTGCAGGGGAGCTCTGGCGCG 270
Db 66 -----MetAlaTyrMetAsp-----Ala 71

QY 271 GCCTACTACCTGGAGATGCTGATCGGGGACCCCGCCGACAGACTACAGATTCTCGTTGAC 330
Db 72 AlaTyrPheGlyGluIleSerIleGlyThrProGlnAsnPheLeuValLeuPheAsp 91

QY 331 ACTGGAAGCAGTAACCTTTGCCGTG-----GCAGGAAGACCCCG 366
Db 92 ThrGlySerSerAsnLeuTrpValProSerValTyrCysGlnSerGlnAlaCysThrSer 111

QY 367 CACTCCTACATAGACAGCTACTTTGACACAGAGAGCTTAGCACATACCGCTCCAGGGC 426
Db 112 HisSer-----ArgPheAsnProSerGluSerSerThrTyrSerThrAsnGly 127

QY 427 TTTCAGGTCACAGTGAAGTACACACAAGGAAGCTGCAGCGGCTTCGTTGGGAAGACCTC 486
Db 128 GlnThrPheSerLeuGlnTyrGlySerGlySerLeuThrGlyPhePheGlyIyrAspThr 147

QY 487 GTCACCATCCCAAGGCTTCAATACTCTTTCTTGTGCAACATTCGCCACTATTTTGAA 546
Db 148 LeuThrVal-----GlnSerIleGlnValProAsnGlnGluPheGlyLeu 162

QY 547 TCAGAGAATTTCTTTTGCCTGG-----ATTAAATGCAATGGAACTACTT 591
Db 163 SerGluAsn-----GluProGlyThrAsnPheValTyrAlaGlnPheAspGlyIleMet 180

QY 592 GCGCTAGCTTATGCCACACTTGGCAAGCCACTCAAGTCTCTCTGGAGACCTTCTCGACTCC 651
Db 181 GlyLeuAlaTyrProAlaLeuSerValAspGluAlaThr-----ThrAlaMetGlnGly 198

QY 652 CTGGTGACAAAGCAACATCCCCAC---GTTTCTCCATCCAGATGTGTGGAGCCGGC 708
Db 199 MetValGlnGluGlyAlaLeuThrSerProValPheSerValTyrLeuSerAsnGln--- 217

QY 709 TTGCCCGTGTGGATCTGGACACACGAGGTAGTGTCTTGGTGGGAATGGAACCA 768
Db 218 -----GlnGlySerSerGlyGlyAlaValvalPheGlyGlyValAspSer 232

QY 769 AGTTTGTTATAAGGAGACATCTCGTATACCCCTATTAAAGAAAGAGTGTACTACAGATA 828
Db 233 SerLeuTyrThrGlyGlnIleTyrTrpAlaProValThrGlnGluLeuTyrTrpGlnIle 252

QY 829 GAAATCTGAAATTGGAATTTGGAGGCCAAAGCCTTAATCTGACTGTCAGAGAGTATAAC 888
Db 253 GlyIleGluGluPheLeuIleGlyGlyGlnAlaSerGlyTrp---CysSerGlu----- 269

QY 889 GCAGACAAGGCCATCTGGACAGTGGCACCCACGCTCTGCGCTGCCCCAGAGGTGTTT 948
Db 270 GlyCysGlnAlaIleValAspThrGlyThrSerLeuLeuThrValProGlnGlnTyrMet 289

QY 949 GATCGGTGGTGGAGCTGTGGCCCGGCACTCTCTGATTCAGAAATTCCTCTGATGGTTTC 1008
Db 290 SerAlaLeuLeuGlnAla----- 295

1009 TGGACTGGGTCCCAAGCTGGCGTGCTCGGACGAATTCGGAACAACCTTGGTCTTACTTC--- 1065

```

```

Db 296 -----ThrGlyAlaGln-----||| ||| ||| ||| |||
2Y 1066 -----CCTAAATCTCCATCTACCTGAGAGATGAGAAC 1098
Db 308 ValAsnCysAsnSerIleGlnAsnLeuProSerLeuThrPheIle-----323
2Y 1099 TCACAGAGTCATTCGGTAICACATCTCGCTAGCTTACATTCACCCCATGAGGG 1138
Db 324 -----AsnGlyValGluPheProLeuProProSerSerfyrIle-----336
2Y 1159 GCCGGCTCAATATCAATGTTAC---CGATTGGCATTCCCA-----TCC 1203
Db 337 -----LeuSerAsnAsnGlyTyrCysThrValGlyValGluProThrTyrLeuSerSer 354
2Y 1204 ACAAAATGCG-----CTGGTATCGTGCCACGGTATGGAGGGCTTCTAGTCATC 1254
Db 355 GlnAsnGlyGlnProLeuTrpIleLeuGlyAspValPheLeuArgSerTyrTyrSerVal 374
2Y 1255 TTCGACAGAGCCCAAGAGGGTGGCTTCGACGGAGC 1293
Db 375 TyrAspLeuGlyAsnAsnArgValGlyPheAlaThrAla 387

RESULT 7
JC7246
pepsinogen C - common marmoset
C:Species: Callithrix jacchus (common marmoset)
C>Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 21-Jul-2000
C:Accession: JC7246
R:Kageyama, T.
J. Biochem. 127, 761-770, 2000
A:Title: New world monkey pepsinogens A and C, and prochymosins. Purification, character
A:Reference number: JC7245
A:Accession: JC7246
A:Molecule type: mRNA
A:Residues: 1-388 <KAG>
A:Cross-references: DDBJ:AB038385
A:Experimental source: strain Nw791
C:Comment: This protein, a zymogen of pepsins, is the major proteolytic enzyme in verte
volution of mammalian orders and families.
C:Superfamily: pepsin
C:Keywords: gastric juice; zymogen

Alignment Scores:
Pred. No.: 2,82e-17 Length: 368
Score: 351.50 Matches: 312
Percent Similarity: 45.16% Conservative: 56
Best Local Similarity: 30.11% Mismatches: 115
Query Match: 10.53% Indels: 89
DB: 2 Gaps: 17

US-09-806-194-1 (1-1804) x JC7246 (1-388)
QY 274 TACTACCTGGATGCTGTATCGGGACCCCGGACAGAGCTACAGATTCTCGTTGACACT 353
Db 73 TyrPheGlyGluIleSerIleGlyThrProGlnAsn:PheLeuValLeuPheAspThr 92
QY 334 GAAGCAGTAACCTTTGCCCTG-----GCAGGAACCCGAC 369
Db 93 GlySerSerAsnLeuTrpValProSerValTyrCysGlnSerGlnAlaCysThrSerHis 112
QY 370 TCTACATAGACAGCTACTTTCACACAGAGAGCTCTAGCACATACCGCTCCAAAGGGCTT 429
Db 113 Ser-----ArgPheAsnProSerAlaSerSerThrTyrSerSerAsnGlyGln 128
QY 430 GAGCTCACAGTCAAGTACACAAAGAGAGCTGGAGGGCTCGTTGGGGAGACACCTTC 489
Db 129 ThrPheSerLeuGlnTyrGlySerGlySerLeuThrGlyPhePheGlyTyrAspThrLeu 148
QY 490 ACCATCCCCAAAGGCTTCAATCTCTTTCTTGTCAACATTCGCCACTATTTTGAATCA 549
Db 149 ThrVal-----GlnSerIleGlnValProAsnGlnGluPheGlyLeuSer 163

```

```

QY 550 GAGAAATCTTTTGTGGTGG-----ATTAAATGGAATGAATACTTGGC 594
Db 164 GluAsn-----GluProGlyThrAsnPheValTyrAlaGlnPheAspGlyIleMetGly 181
QY 595 CTAGCTTATCCACACTGCGAAGCCATCAAGTCTCTCGAGACCTTCTTCGACTCCCTG 654
Db 182 LeuAlaTyrProAlaLeuSerMetGlyGlyAlaThr-----ThrAlaMetGlnGlyMet 199
QY 655 GTACACAAACAAACATCCCAAC-----GTITTCCTCATCCAGATGTGTGGAGCCGCTT 711
Db 200 LeuGlnGluGlyAlaLeuThrSerProValPheSerPheTyrLeuSerAsnGln-----217
QY 712 CCGTGTCTGGATCTGGACCAACGAGGTAGTCTTGTCTGGTGGTGAATGAACCAAGT 771
Db 218 -----GlnGlySerSerGlyGlyAlaValIlePheGlyGlyValAspSerSer 233
QY 772 TTGTATAAGGACACATCTGTATACCCCTATTAAAGAGAGAGTGGTACTACAGATAGAA 831
Db 234 LeuTyrThrGlyGlnIleTyrTrpAlaProValThrGlnGluLeuTyrTrpGlnIleGly 253
QY 832 ATTCTGAATTTGAAATTGGAGCCCAAGCCTTAATCTGGACTGACAGAGATATAACGA 891
Db 254 IleGluGluPheLeuIleGlyGlyGlnAlaSerGlyTrp---CysSerGlu-----Gly 270
QY 892 GACAAGGCCATCTGGACACTGGCACCGCTGCTGGCTGCCCCAGAGGTGTTGAT 951
Db 271 CysGlnAlaIleValAspThrGlyThrSerLeuLeuThrValProGlnGlnTyrMetSer 290
QY 952 CGGTGGTGGAGCTGTGGCCCGCATCTCTGATTCAGAAATCTCTGTGTTCTG 1011
Db 291 AlaPheLeuGluAla-----295
QY 1012 ACTGGTCCAGCTGGCGTCTCGACGAATTCGGAACACCTTGTCTTACTTC-----1065
Db 296 ThrGlyAlaGln-----GluAspGluTyrGlyGlnPheLeuVal 308
QY 1066 -----CCTAAATCTCCATCTACCTGACGAGATGAGAACTCC 1101
Db 309 AsnCysAspSerIleGlnAsn:LeuProThrLeuThrPheIleIle-----323
QY 1102 AGCAGGTCACTCCGTATCAAAATCCCTGCTCAGCTTTTACATTCAGCCCATGATGGGGCC 1161
Db 324 ---AsnGlyValGluPheProLeuProProSerSerfyrIle-----336
QY 1162 GGCTGAATATGAATGTTAC---CGATTGGCATTTCCCA-----1200
Db 337 ---LeuSerAsnAsnGlyTyrCysThrValGlyValGluProThrTyrLeuSerSerGln 355
QY 1201 ---TCCACAAATCGCTGGTGTGTCGTCGCCACGGTATGGAGGGCTTCTACGTACTTC 1257
Db 356 AsnSerGlnProLeuTrpIleLeuGlyAspValPheLeuArgSerTyrTyrSerValPhe 375
QY 1258 GACAGACCCCAAGAGGGTGGCTTCGACGCGAGC 1293
Db 376 AspLeuGlyAsnAsnArgValGlyPheAlaThrAla 387

RESULT 8
B43356
gastricsin (EC 3.4.23.3) precursor - guinea pig
R:Alternate names: pepsin C
C:Species: Cavia porcellus (guinea pig)
C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 22-Jun-1999
C:Accession: B43356
R:Kageyama, T.; Ichinose, M.; Tsukada, S.; Miki, K.; Kurokawa, K.; Kolwai, O.; Tanji,
J. Biol. Chem. 267, 16450-16459, 1992
A:Title: Gastric procathepsin E and progastricsin from guinea pig. Purification, mole
A:Reference number: A43356; MUID:92355614; PMID:1644829
A:Accession: B43356
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-394 <KAG>
A:Cross-references: GB:M88652; NID:gl91296; PIDN:AAA37053.1; PID:gl91297
A>Note: sequence extracted from NCBI backbone (NCBIN:110805, NCBIP:110806)

```


Superfamily: pepsin
 Keywords: aspartic proteinase; gastric juice; hydrolase; protein digestion; stomach

Alignment Scores:
 Score: 2.45e-15 Length: 394
 Percent Similarity: 324.50 Matches: 107
 Best Local Similarity: 46.07% Conservative: 63
 Query Match: 29.00% Mismatches: 116
 Indels: 83
 Gaps: 18

S-09-806-194-1 (1-1804) x B43356 (1-394)

Y 274 TACTACCTGGAGATCTGTGATCGGACCCCGCCGAGAGCTACAGATTCTCGTGTGACACT 333
 b 79 TyrPheGlyGlnIleSerLeuGlyThrProGlnSerPheGlnValLeuPheAspThr 98
 Y 334 GGAGCAGTAACCTTT-----GCCGTGGCAGGAACCCGAC 369
 b 99 GlySerSerAsnLeuTrpValProSerValTyrCysSerLeuAlaCysIleThrHis 118
 Y 370 TCCATACATAGACAGCTACTTTGACACAGAGGCTAGCACATACCGCTCCAAGGCGT 429
 b 119 -----ThrArgPheAsnProArgAspSerSerThrIleValAlaIleAspGln 134
 Y 430 GACGTCACAGTGAAGTACACAAAGAGCTGGACGGCTTCGTTGGGGAAGACCTGGTC 489
 b 135 SerPheSerLeuGluTyrGlyThrGlySerLeuThrGlyValPheGlyTyrAspThrMet 154
 Y 490 ACCATC-----CCCAAA--CGCTTCAATACTTCITTTCTGTGACACT 531
 b 155 ThrIleGlnAspIleGlnValProGlyGlnLysPheGlyLeuSer----- 169
 Y 532 GCCACTATTTTGAATCAGAGAAATTTCTTTTGCTGGG-----ATAAA 576
 b 170 -----GluThrGlu-----ProGlySerAspPheValIleAlaGlu 181
 Y 577 TGGAAATGAATACTTGGCTAGCTTTATCCACACTTCCAGACCATCAAGTCTCTGGAG 636
 b 182 PheAspGlyIleLeuGlyLeuGlyTyrProGlyLeuSerGluGlyGlyAlaThr----- 199
 Y 637 ACCTTCTTCGACTCCCTGGTGACACAAAGAACATC--CCCAAGCTTTCTCATCGAC 693
 b 200 ThrAlaMetGlnGlyLeuLeuArgGluGlyAlaLeuSerGlnSerLeuPheSerValTyr 219
 Y 694 ATGTGTGAGCGCGCTTCCGCTTGATCT-----GGGACCAAGGAGTAGTCTT 747
 b 220 Leu-----GlySerGlnGlnGlySerAspGluGlyGlnLeu 231
 Y 748 GTCTTGGTGAATTAACCAAGTTTGATTAAGGACACATCTGTATACCCCTATTAG 807
 b 232 IleLeuGlyGlyValAspLysLeuThrGlyAspIleTyrTrpThrProValThr 251
 Y 808 GAAGACTGGTACTACACAGATAAATCTGAAATTTGGAATTTGGAGCCAAAGCTTAAT 867
 b 252 GlnGluLeuTyrTrpGlnIleGlyIleGluGlyPheLeuIleAspGlySerAlaSerGly 271
 Y 868 CTGGACTCCAGAGATATAAGCCAGACAGCCCATCGTGACATGGCACACCGCTCGT 927
 b 272 TrpCysSerArg-----GlyCysGlnGlyIleValAspThrGlyThrSerLeuLeu 288
 Y 928 CGCTGCCCCAGAGGTGTTGATCGGTGGTGAAGCTGTGGCCCGCGCATCTCTGATT 987
 b 289 ThrValProSerAspTyrLeuSerThrLeuValGlnAlaIleGlyAlaGluGlu----- 306
 Y 988 CCAAGATTCTGTGGTGTTCCTGGACTGGTCCGACATGGCTGGCTGGCAAGTTCGGAA 1047
 b 307 AsnGluTyrGlyGluTyrPhe-----ValSerCysSerSerIleGluAsp 321
 Y 1048 ACACCTTGGTCTTACTCCCTAAATCTCCATCTACCTGAGAGATGAGAACTCCAGCG 1107
 b 322 LeuProThrLeuThrPheValIleSerGlyVal----- 332

QY 1108 TCATTCCTGATACAAATCTGCTCAGCTTACATTCAGGCC-----ATG 1152
 Db 333 -----GluPheProLeuSerProSerAlaTyrIleLeuSerGlyGluAsnTyrCysMet 350
 QY 1153 ATGGGGCGCGCTGAATTAATGTTTACCGATTTCGGCATTTCCCATCCACAAT--- 1209
 Db 351 ValGlyLeuGluSerThrTyr-----ValSerProGlyGlyGlyGlu 364
 QY 1210 ---GGCTGGTATCGGTGCCACGGTGTATGGAGGCTTCTACGTATCTTCGACAGAGCC 1266
 Db 365 ProValTrpIleLeuGlyAspValPheLeuArgSerTyrIleValTyrAspLeuAla 384
 QY 1267 CAGAAGAGGGTGGCTTCGACGAGC 1293
 Db 385 AsnAsnArgValGlyPheAlaThrAla 393

RESULT 9
 JC7575
 pepsinogen A - bullfrog
 C:Species: Rana catesbeiana (bullfrog)
 C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 03-Aug-2001
 A:Accession: JC7575
 K:Ikuzawa, M.; Inokuchi, T.; Kobayashi, K.; Yasumasu, S.
 J. Biochem. 129, 147-153, 2001
 A:Title: Amphibian pepsinogens: Purification and characterization of Xenopus pepsinog
 A:Reference number: JC7573; MUID:21064922; PMID:11134969
 A:Contents: Stomach
 A:Accession: JC7575
 A:Molecule type: mRNA
 A:Residues: 1-385 <LKU>
 A:Cross-references: DDBJ:AB045376
 C:Comment: This protein is a zymogen for gastric aspartic proteinase, with pepsin-lik
 C:Genetics:
 A:Gene: Pga
 C:Superfamily: pepsin
 C:Keywords: stomach; zymogen

Alignment Scores:
 Pred. No.: 5,114e-15 Length: 385
 Score: 320.00 Matches: 121
 Percent Similarity: 44.61% Conservative: 67
 Best Local Similarity: 27.82% Mismatches: 147
 Query Match: 9.58% Indels: 74
 Gaps: 2

US-09-806-194-1 (1-1804) x JC7575 (1-385)

QY 148 GGGACCCCTGGCGAGGCGCACCGCCGCGCTTGGCGCTCGCCCTGGAGCCTGCCCTGGCG 207
 Db 339 GlyAspTyrLeuLysLysHisIleTyrAsnProAlaThrLysTyrPheProSerLeuAla 58
 QY 208 TCCCGCGGCGCGCGCCCAACTTCTTGGCCATGTGTAGACAACCTGCGAGGGGACTCTGGC 267
 Db 59 GlnAlaSerGly-----GluProLeuGlnAsnTyrMetAsp 70
 QY 268 CCGCGCTACTACTCGAGATGCTGATCGGAGCCCGCCGAGAGCTACAGATTCTCGTT 327
 Db 71 IleGluTyrPheGlyThrIleSerIleGlyThrProGlnSerPheThrValIlePhe 90
 QY 328 GACACTGGAAGCAGTAACCTTTGCCGTGGCAGGAACCCCGCACTCTACATAGACACG--- 384
 Db 91 AspThrGlySerSerAsnLeuTrpVal-----ProSerValTyrCysSerPro 107
 QY 385 -----TACTTTCACAGAGAGGCTACACATACCCGCTCCACGGC 426
 Db 108 AlaCysThrAsnHisHisMetPheAsnProGlnSerSerThrPheGlnAlaThrAsn 127
 QY 427 TTTGAGCTCACAGTGAAGTACACACAAGGAGCTGGAGGCTTCGTGGGGAAGACCTC 486
 Db 128 ThrProValSerIleGlnIleGlyThrGlySerMetSerGlyPheLeuGlyTyrAspThr 147
 QY 487 GTCAACATCCCCAAAGGCTTCAATCTTCTTTCTGTCAACATTCGCCACTATTTTTAA 546

520 CTGTCACATGCCACTATTTCGAATCAGACAAATTCCTT-----TTCCT 567
 166 -----LeuValValValAspGlnGluPheIleGluAlaThrLysGluPro 180
 568 GGGATT-----AAATGGAAATGAATACTTGGCTAGCTTATGCCACAT 612
 181 GlyLeuThrPheMetValAlaLysPheAspGlyIleLeuGlyLeuGlyPheGlnGluIle 200
 613 GCCAAGCCATCAAGTCTCTGGAGACCTTCTCGACTCCCTGGTGACACAGCAACAATC 672
 201 SerValGlyAspAla-----Val 206
 673 CCCAAGCTTCTCCATCCAGATGCTGGA---GCCGGCTTCCCGTTCCTGGATCGG 729
 207 ProValTrpTyrLysMetValGluGlnGlyLeuValSerGluProValPheSerPheTrp 226
 730 ACCAAC-----GGAGTAGTCTTGTCTTGGTGGAAATGAACA 768
 227 PheAsnArgHisSerAspGluGlyGluGlyGluIleValPheGlyGlyMetAspPro 246
 769 ACTTCTTATAAGGACATCTCGTATACCCCTATTAAAGGAAGACTGCTACTACCAATA 828
 247 SerHisTyrLysGlyAsnHisThrTyrValProValSerGlnLysGlyTyrTrpGlnPhe 266
 829 GAAATCTGAAATTCGAATTCGAGCAAGCAAGCCCTTAATCTCGACTGCAGAGATATAAC 888
 267 GluMetGlyAspValLeuIleGlyLysThrThrGlyPhe---CysAla-----Ser 283
 889 CGACACAAGGCCATCGTGACACTGGCACCAGCGTCTGGCGCTGCCCGCAAGAGTGT 945
 284 GlyCysSerAlaIleAlaAspSerGlyThrSerLeuLeuAlaGlyProThrAlaIle 303
 949 GATCGGTGGTGAAGCTGTGGCCCGCGCATCTCTGATTCAGAA----- 993
 304 ThrGluIleAsnGluLysIleGlyAlaThrGlyValValSerGlnGluCysLysThrVal 323
 993 ----- 993
 324 ValSerGlnTyrGlyGlnGlnIleLeuAspLeuLeuAlaGluThrGlnProSerLys 343
 993 ----- 993
 344 IleCysSerGlnValCysLysThrPheAspGlyLysHisGlyValSerAlaGlyIle 363
 994 -----TTCCTGATGCTTC----- 1008
 364 LysSerValValAspAspGluAlaGlyGluSerAsnGlyLeuGlnSerGlyProMetCys 383
 1009 -----TGGACTGGTCCCGACGTCGCTGGACGCAATTCG 1044
 384 AsnAlaCysGluMetAlaValValTrpMetGlnAsnGlnLeuAlaGlnAsnLysThrGln 403
 1045 GAACACCTTGCTCTTAC----- 1062
 404 AspLeuIleLeuAsnTyrIleAsnGlnLeuCysAspLysLeuProSerProMetGlyGlu 423
 1063 -----TCCCTAAATCTCATCTACCTGAGACAT 1092
 424 SerSerValAspCysGlySerLeuAlaSerMetProGlnIleSerPheThrIleGlyAla 443
 1093 GAGAACTCCAGCAGGTCAATCCGTATCAACAATCTCGCTACCACTTTACATTCAGCCCATG 1152
 444 Lys-----LysPheAlaLeuLysProGluGluTyrIleLeu---Lys 456
 1153 ATGGGGCGCGCTGAATATGAATGTTAC-----CGATCGGCTATTCGCA 1200
 457 ValGlyGluGlyAlaAlaGlnCysIleSerGlyPheThrAlaMetAspIleProPro 476
 1201 TCCACAATGCGCTG---GTGATCGTCCCGGATGGGGCTTCTACCTCATCTTC 1257
 477 ProArgGlyProLeuTrpIleLeuGlyAspValPheMetGlyAlaTyrHisThrValPhe 496

QY 1258 GACAGAGCCAGAGAGGGTGGCTTGGCAGCGAGC 1293
 Db 497 AspTyrGlyLysMetArgValGlyPheAlaLysSer 508
 RESULT 12
 A24608
 Gastricisin (EC 3.4.23.3) precursor - rat
 N:Alternate names: pepsinogen C
 N:Contains: pepsin A (EC 3.4.23.1) precursor
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 30-Jun-1988 #sequence,revision 05-Aug-1994 #text_change 18-Jun-1999
 C:Accession: A33510; A24608; C22434; A05145; A61298
 R:Ichihara, J.; Ichihara, Y.; Hayano, T.; Katsura, I.; Sogawa, K.; Fujii-Kuriyama, Y.
 J. Biol. Chem. 264, 10193-10199, 1989
 A:Title: Primary structure and transcriptional regulation of rat pepsinogen C gene.
 A:Reference number: A33510; MUID:89235508; PMID:2722863
 A:Molecule type: DNA
 A:Accession: A33510
 A:Residues: 1-392 <ISH>
 A:Cross-references: GB:M25985
 R:Ichihara, Y.; Sogawa, K.; Morohashi, K.; Fujii-Kuriyama, Y.; Takahashi, K.
 Eur. J. Biochem. 161, 7-12, 1986
 A:Title: Nucleotide sequence of a nearly full-length cDNA coding for pepsinogen of ra
 A:Reference number: A24608; MUID:87054020; PMID:3780741
 A:Accession: A24608
 A:Molecule type: mRNA
 A:Residues: 1-392 <ICH>
 A:Cross-references: GB:X04644; NID:956880; PIDN:CAA28305.1; PID:956881
 R:Ichihara, Y.; Sogawa, K.; Takahashi, K.
 J. Biochem. 98, 483-492, 1985
 A:Title: Isolation of human, swine, and rat prepepsinogens and calf preprochymosin, a
 A:Reference number: A22434; MUID:86059312; PMID:2415509
 A:Accession: C22434
 A:Molecule type: protein
 A:Residues: 1-19, X', 21-23, X', 25-29 <IC2>
 R:Arai, K.M.; Muto, N.; Tani, S.; Akahane, K.
 Biochim. Biophys. Acta 788, 256-261, 1984
 A:Title: The N-terminal sequence of rat pepsinogen.
 A:Reference number: A05145; MUID:84257697; PMID:6743670
 A:Accession: A05145
 A:Molecule type: protein
 A:Residues: 17-30, Q', 32-102, A', 104-108, L', 110-112 <ARA>
 R:Ichihara, Y.; Sogawa, K.; Takahashi, K.
 J. Biochem. 92, 603-606, 1982
 A:Title: Rat gastric propepsinogen: in vitro synthesis and partial amino-terminal sig
 A:Reference number: A61298; MUID:83030750; PMID:6182139
 A:Accession: A61298
 A:Molecule type: protein
 A:Residues: 1, 'XX', 4-6, 'X', 8-9, 'X', 11, 'X', 13-14, 'XXX', 18-19, 'X', 21, 'X', 23, 'XX', 25, 'X'
 C:Comment: This enzyme has more restricted specificity than pepsin A. It is the major
 C:Genetics:
 A:Introns: 20/2; 73/3; 113/1; 152/3; 219/2; 259/2; 309/3; 342/3
 A:Note: there are at least two very similar genes for gastricsin in rat
 C:Superfamily: pepsin
 C:Keywords: aspartic proteinase; gastric juice; hydrolase; protein digestion; stomach
 F:1-16/Domain: signal sequence #status experimental <SIG>
 F:17-392/Product: pepsinogen #status experimental <MAT>
 F:17-62/Domain: activation peptide #status experimental <ACT>
 F:94, 280/Active site: Asp #status predicted
 F:107-112, 270-275, 314-347/Disulfide bonds: #status predicted

Alignment Scores:
 Pred. No.: 1.64e-14 Length: 392
 Score: 313.00 Matches: 105
 Percent Similarity: 45.22% Conservative: 56
 Best Local Similarity: 29.49% Mismatches: 139
 Query Match: 9.37% Indels: 56
 DB: 1 Gaps: 16

US-09-806-194-1 (1-1804) x A24608 (1-392)

QY 274 TACTACCTGGAGATGCTGATGGGACCCCGGACAGCTACAGATTCCTGTTGACACT 333

76 TyrPheGlyGluIleSerIleClyThrProGlnAsnPheLeuValLeuPheAspThr 95
 334 GAAAGCAGTAACCTTTGCCGTG-----GAGGAAACCCCGCCAC 369
 96 GlySerSerAsnLeuIrpValSerSerValTyrCysGlnSorGluAlaCysThrHis 115
 370 TCCTACATAGACAGCTACTTTCACACAGAGAGTCTACGACATACCGCTCCAAAGGGCTTT 429
 116 Ala-----ArgPheAsnProSerLysSerSerThrTyrThrGluGlyGln 131
 430 GAGGTCACAGTCAAGTACACACAAGGAAGCTGGAGGGCTTCGTGGGAGAGACTCGTC 489
 132 ThrPheSerLeuGlnTyrClyThrGlySerLeuThrGlyPhePheGlyTyrAspThrLeu 151
 490 ACCATCCCCAAAGGCTTCAATACTCTTTCTTGTGCAACATTCGCCACTATTTTGAATCA 549
 152 ThrVal-----GluSerIleGlnValProAsnGlnGluPheGlyLeuSer 166
 550 GAGAAATTTCTTTTGCCTGGG-----A-TAAATGGAAATGGAAATCTTGGC 594
 167 GluAsn-----GluProGlyThrAsnPheValTyrAlaGlnPheAspGlyIleMetGly 184
 595 CTAGCTTATGGCCACACITGCCAACCCATCAAGTTCTCTGGAGACCTTCTTCGACTCGCTG 654
 185 LeuAlaTyrProGlyLeuSer-----SerGlyGlyAlaThrThrAlaLeuGlnGlyMet 202
 655 GTGACACAGCAACATCCCAACGTTTCTCCATGCAGATGTGTGGAGCGGGTGGCC 714
 203 LeuGlyGlu-----GlyAlaLeuSerGlnProLeuPheGlyValTyrLeu--- 217
 715 GTTCTGGATCT-----GGGACCAACGGAGCTAGTCTGTCTGGTGGAAATGSAACCA 768
 218 -----GlySerGlnGlnGlySerAsnGlyGlyGlnIleValPheGlyGlyValAspLys 235
 769 AGTTGTATAAGGAGACATCTGCTATACCCCTATTANGAGAGAGTGTACTACAGATA 828
 236 AsnLeuThrGlyGluIleThrIrpValProValThrGlnGluLeuGlyTyrGlnIle 255
 829 GAAATCTGAAATTTGGAAATTTGGAGGCCAAAGCCCTTAATCTGGACTCGAGAGATTAAC 885
 256 ThrIleAspAspPheLeuIleGlyAspGlnAlaSerGlyIrp---CysSerSerGlnGly 274
 889 CGACAGAGCCATCGTGGACAGTGGCACACGCTGCTCCGCTGCCCGCCACAGAGTGTTT 948
 275 Cys---GlnGlyIleValAspThrGlyThrSerLeuLeuValMetProAlaGlnTyrLeu 293
 949 GATCGGTGGTGGAGGCTGGCGCGCCATCTCTGATTCAGAGATCTCTGATGATGTTG 1008
 294 SerGluLeuLeuGlnThrIleGlyAlaGlnGlu-----GlyGluTyrGlyGlyTyrPhe 311
 1009 TGGACTGGTCCGAGCTGGCTGTGAGCAGAACTTCGGAACACACCTTGGTCTACTTCGCT 1068
 312 -----ValSerCysAspSerValSer-----LeuPro 321
 1069 AAAATCTCCATCTACCTCAGAGATCAGAACTCCAGCAGGTCAATCCGCTATCACAACTG 1128
 322 ThrLeuSerPheValLeu-----AsnGlyValGlnPheProLeuSer 335
 1129 CCTCAGCTTTAC---ATTCAGCCCATGATGGGGCGCGCCTGAAATATGAATGTTACCGA 1185
 336 ProSerSerTyrIleIleGlnGluAspAsnPheCysMetValGlyLeuGluSerIleSer 355
 1186 ITGGCAATTTCCCATCCCAAAATCGCTGCTGATCGGTGGCCAGCGTATCGAGGGGCTTC 1245
 356 LeuThrSerGluSerGlyGlnProLeuThrIleLeuGlyAspValPheLeuArgSerTyr 375
 1246 TACGCTCATCTTCGACAGAGCCGAAAGAGGGTGGGCTTCGACGAGCG 1293
 376 TyrAlaIlePheAspMetGlyAsnAsnLysValGlyLeuAlaThrSer 391

pepin (EC 3.4.23.-) precursor, embryonic - chicken
C:Species: Gallus gallus (chicken)
C:Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 *text_change 21-Jul-2000
C:Accession: A1443
R:Hayashi, K.; Aizawa, K.; Mochii, M.; Yasugi, S.; Eguchi, G.; Mizuno, T.
J. Biochem. 103, 290-296, 1988
A:Title: Molecular cloning and the nucleotide sequence of cDNA for embryonic chicken
A:Reference number: A1443; MUID:8827903; PMID:3131317
A:Accession: A1443
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-383 <HAY>
A:Cross-references: GB:D00215; NID:g2760810; PIDN:BAA00153.1; PID:g222853
C:Superfamily: pepsin
C:Keywords: aspartic proteinase; hydrolase; protein digestion

Alignment Scores:
Pred. No.: 2,68e-14 Length: 383
Score: 310.00 Matches: 106
Percent Similarity: 42.78% Conservative: 63
Best Local Similarity: 26.84% Mismatches: 136
Query Match: 9.28% Indels: 90
DB: 2 Gaps: 15

US-09-806-194-1 (1-1804) x A1443 (1-383)

Qy 166 CAGGCC-----GACGGCTGGCCCTGCCTCGAGCTGCCTGGCGTCCCTCCCGCGGGC 219
Db 55 H1SAlaPheProaspValLeuThrValValThrGluProLeuLeu----- 69
Qy 220 GCGGCCAAACTCTTGGCCATGGTAGACAACCTGCAGGGGGACTCTGCCGGCGGTACTAC 279
Db 70 -----AsnThrLeuAspMet-----GluTyrTyr 77
Qy 280 CTGAGAGTCGTATCGGAGCCCGCCCGCAGAGGTACAGATTCTCGTTGACACTGGAAGC 339
Db 78 GlyThrIleSerIleGlyThrProGlnAspPheThrValValPheAspThrGlySer 97
Qy 340 AGTAACTTTGCCGTGGCAGGA-----ACCCCGCAGCTCCTACATAGACACGTAC 387
Db 98 SerAsnLeuThrValProSerValSerCysThrSerProAlaCysGlnSerHisGlnMet 117
Qy 388 TTTCACACAGAGGCTTAGCACATACCGCTCCAGGGCTTTGACGTACAGTCAAGTAC 447
Db 118 PheAsnProSerGlnSerThrTyrIleSerThrGlyGlnAsnLeuSerIleHisTyr 137
Qy 448 ACACAAGGAAGCTGGACGGGCTTCGTTGGGAAGACCTCGTCACCATCCCCAAGGGCTC 507
Db 138 GlyThrGlyAspMetGluGlyThrValGlyCysAspThrValThrValAlaSerLeuMet 157
Qy 508 AATACTCTTTCTTGCACATAGCCCATATTTTGAATCAGAGAATTCCTTTTGGCT 567
Db 158 AspThrAsnGlnLeuPheGlyLeuSerThr---SerGluProGlyGlnPhePheVal--- 175
Qy 568 GGGATTAAATGGAAATGAAATCTATGGCTAGCTTAGTCACACTTGCACAGCCCAAGTCAAGT 627
Db 176 TyrValIysPheAspGlyIleLeuGlyLeuGlyTyrProSerLeuAla-----AlaAsp 193
Qy 628 TCTCTGGAGACTTCTTCGACTCCCTGGTGACACAGCAACATC---CCCAAGCTTTTC 684
Db 194 GlyIleThrProValPheAspAsnMetValAsnGluSerLeuLeuGluGlnAsnLeuPhe 213
Qy 685 TCCATGCAGATGTGTGAGCGGGCTTCCCGTCTGCTGATCTGGACCAAGGAGGTAGT 744
Db 214 SerValTyrLeuSerArgGluProMet-----GlySerMet 225
Qy 745 CTGTCTGTGGTGGAAATGAACCAAGTTTGTATAAAGGAGACATCTGGTATACCCCTATT 804
Db 226 ValValPheGlyGlyIleAspGluSerTyrPheThrGlySerIleAsnTrpIleProVal 245
Qy 805 AAGGAAGAGTGGTACTACCAAGATAGAAATCTGAAATTTGGAATTTGGAGGCCCAAGCCTT 864
Db 246 SerTyrGlnGlyTyrTrpGlnIleSerMetAspSerIleIleValAsnLysGlnGluIle 265


```
QY 547 TCAGAGAAATTTCTTTTCCTGGGATT-----AATGGAATCGAATACCT 591
Db      |||
178 glyGluAlaThrLysGlnProGlyTleValPheValAlaAlaLysPheAspGlyLeu 197
QY 592 GGCCTAGCTTATGCCACACACTTCCCAAGCCATCAAGTCTCTGGAGACCTTCTTGACTCC 651
Db      |||
198 GlyMetGlyTyrProHisIleSer-----ValAsnAsnValLeuProValPheAspAsn 215
QY 652 CTGGTGACACAGCAACAATC---CCCAACGTTTCTCCATCGACATGTGGAGCGGCG 708
Db      |||
216 LeuMetGlnGlnLysLeuValAspLysAsnIlePheSerPheTyr----- 230
QY 709 TTGCCCGTTGCTGGATCTGGACCAACGAGGAGTAGTCTTGTTGGGTGGAAATGCAACCA 768
Db      |||
231 LeuAsnArgAspProGluGlyGlnProGlyGlyGluLeuMetLeuGlyGlyThrAspSer 250
QY 769 ACTTTGTATAAGAGGAGACATCTGGTATACCCCTATTAAAGGAAGAGTGGTACTACAGATA 828
Db      |||
251 LysTyrTyrHisGlyGluLeuSerTyrLeuAsnValThrArgLysAlaTyrTyrPoleVal 270
QY 829 GAAATTCGTAATTCGAATTCGAGGCGCAAGCCCTTAATCTCGACTGCAGACAGATATAAC 888
Db      |||
271 HisMetAspGlnLeuGluValGlyAsnGlu---LeuThrLeu---CysLys-----Gly 286
QY 889 CGACACAAGGCCATCGTCGACAGTGGCAGACCGACCGCTGCTGGCGCTGCCCGCAGAGAGTGT 948
Db      |||
287 GlyCysGluAlaIleValAspThrGlyThrSerLeuLeuValGlyProValGluGluVal 306
QY 949 GATCGGTGGTGAAGCTGTGGCCCGCGCATCTCTGATTCAGATTCACAGAATCTCTGATGGTTTC 1008
Db      |||
307 LysGluLeuGlnLysAlaIleGlyAlaValProLeuIle----- 319
QY 1009 TGGACTGGGTCCCACTGGCGTGTGGAGCAATTCGGAACACCTTGGTCTTACTTCCT 1069
Db      |||
QY 1069 ARAATCTCATCTACCTGAGAGATGAGAACTCCACAGGTCAATCCGTATCACAATCCTG 1128
Db      |||
320 -----GlnGlyGluTyrMetIleProCysGluLysValSerSerLeu 333
QY 1129 CCTCAGCTTACATTCAGCCCATGATGGGGCGCGCTGAATTATGAATGT----- 1179
Db      |||
334 ProThrValTyrLeuLys-----LeuGly---GlyLysAsnTyrGluLeuHisProAsp 350
QY 1180 ---TACCGATTCCGCATTTC----- 1197
Db      |||
351 LysTyrIleLeuLysValSerGlnGlyLysThrIleCysLeuSerGlyPheMetGly 370
QY 1198 -----CCATCCACAAATGGCTGGTGCATCGGTGCCACGGTATGGAGGCG 1242
Db      |||
371 MetAspIleProProSerGlyProLeuThrIleLeuGlyAspValPheIleLys 390
QY 1243 TTCTAGCTCATCTTCACAGACCGCAGAGAGGGTGGCTTCGCA 1287
Db      |||
391 TyrTyrThrValPheAspArgAsnAsnArgValGlyPheAla 405
```

Search completed: September 25, 2003. 11:09:36
Job time : 91 secs

Result No.	Score	Query Match	Length	DB ID	Description
1	100	100	100	1	...
2	95	95	95	2	...
3	90	90	90	3	...
4	85	85	85	4	...
5	80	80	80	5	...
6	75	75	75	6	...
7	70	70	70	7	...
8	65	65	65	8	...
9	60	60	60	9	...
10	55	55	55	10	...
11	50	50	50	11	...
12	45	45	45	12	...
13	40	40	40	13	...
14	35	35	35	14	...
15	30	30	30	15	...
16	25	25	25	16	...
17	20	20	20	17	...
18	15	15	15	18	...
19	10	10	10	19	...
20	5	5	5	20	...

.....

; PRIOR FILING DATE: 1999-09-23

: PRIOR APPLICATION NUMBER: 80/133,499
 : PRIOR FILING DATE: 1999-09-23

; PRIOR FILING DATE: 1999-09-23

Db	241	GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrIys	260
QY	781	GGACACATCTGGTATACCCCTATTAGGAGAGACTGCTACTCCACGATAGCAATTCCTGAAA	840
Db	261	GlyAspIleThrProIleLysGluGluTrpIleGluIleLeuIys	280
QY	841	TTGGAATTGGAGCCCAAGCCCTTAATCTGGACTGCAGAGATATAACGCACAGAACGCC	900
Db	281	LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla	300
QY	901	ATCTGGACACTGCCACACGCTGCTGCCCTGCCCCACAAAGGTGTTTGATGCCGGTGTG	960
Db	301	IleValAspSerGlyThrLeuLeuArgLeuProGlnLysValPheAspAlaVal	320
QY	961	GAAGCTGTGGCCCGCATCTCTCATTCAGAAATCTCTGATGGTTCCTGGACTGGGTCC	1020
Db	321	GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer	340
QY	1021	CAGCTGGCGTCTGGAGCAATTCGGAAACACCTTGGTCTTACTTCCTTAAATCTCCATC	1080
Db	341	GlnLeuAlaCysIleThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle	360
QY	1081	TACCTGACAGATGAGAAGCTCCAGCAGGTCATTCGGTATCACAAATCTCGCTCAGCTTAC	1140
Db	361	TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr	380
QY	1141	ATTCAGCCCATGTATGGGGGGCGGCTGAATTATGAATGTTACCGATTTCGCATTTCGCCA	1200
Db	381	IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro	400
QY	1201	TCCCAAAATCGCTGGTGATCGGTGGCCAGCGTGTAGGAGGCTTCTACGTCATCTTCGAC	1260
Db	401	SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheasp	420
QY	1261	AGAGCCCAAGAGGGTGGGCTTCGACGACGAGCCCTGTGCGAAGAAATGCAGGTCCTGCA	1320
Db	421	ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla	440
QY	1321	GTGCTGAAATTTCCGGGCGCTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCGCCGCT	1380
Db	441	ValSerGluIleSerGlyProPheSerThrGluaspValAlaSerAsnCysValProAla	460
QY	1381	CAGTCTTTGACGAGCCCATTTTGTGGATTGTCTCTATGCGCTCATGAGCGTGTGTGGA	1440
Db	461	GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly	480
QY	1441	GCCATCTCTTGTCTTAATCGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1500
Db	481	AlaIleLeuLeuValIleValLeuLeuLeuLeuProPheArgCysGlnArgArgPro	500
QY	1501	COTGACCTGAGGTGCGCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA	1554
Db	501	ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys	518
RESULT 4			
CS-09-794-748-2			
: Sequence 2, Application US/09794748			
: Patent No. US20020037315A1			
: GENERAL INFORMATION:			
: APPLICANT: Gurney, Mark E.			
: APPLICANT: Bienkowski, Michael J.			
: APPLICANT: Heinrichson, Robert L.			
: APPLICANT: Parodi, Luis A.			
: APPLICANT: Yan, Riqiang			
: TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR,			
: TITLE OF INVENTION: USES			
: TITLE OF INVENTION: THEREFOR			
: FILE REFERENCE: 28341/6280JL			
: CURRENT APPLICATION NUMBER: US/09/794,748			
: PRIOR FILING DATE: 2001-02-27			
: PRIOR APPLICATION NUMBER: 09/416,901			
: PRIOR FILING DATE: 1999-10-13			

RESULT 4
C-87-704-728-3

```

CS-09-794-748-2
Sequence 2, Application US/09794748
Patent No. US20020037315A1
GENERAL INFORMATION:
APPLICANT: Gurney, Mark E.
APPLICANT: Blenkowski, Michael J.
APPLICANT: Heinrichson, Robert L.
APPLICANT: Parodi, Luis A.
APPLICANT: Yan, Riqiang
TITLE OF INVENTION: ALZHEIMER'S D
TITLE OF INVENTION: USES
TITLE OF INVENTION: THEREFOR
FILE REFERENCE: 28341/6280JL
CURRENT APPLICATION NUMBER: US/09/
CURRENT FILING DATE: 2001-02-27
PRIOR APPLICATION NUMBER: 09/4716,
PRIOR FILING DATE: 1999-10-13

```


: LENGTH: 518

APPLICANT: Shelton, David L.	?	PRIOR FILING DATE: 1998-04-01
APPLICANT: Stewart, Timothy A.	?	PRIOR APPLICATION NUMBER: 60/081070
APPLICANT: Tumas, Daniel	?	PRIOR FILING DATE: 1998-04-08
APPLICANT: Williams, P. Mickey	?	PRIOR APPLICATION NUMBER: 60/081049
APPLICANT: Wood, William I.	?	PRIOR FILING DATE: 1998-04-08
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic	?	PRIOR APPLICATION NUMBER: 60/081071
TITLE OF INVENTION: Acids Encoding the Same	?	PRIOR FILING DATE: 1998-04-08
FILE REFERENCE: P2630Pic11	?	PRIOR APPLICATION NUMBER: 60/081195
CURRENT APPLICATION NUMBER: US/09/978,295A	?	PRIOR FILING DATE: 1998-04-08
CURRENT FILING DATE: 2001-10-15	?	PRIOR APPLICATION NUMBER: 60/081203
PRIOR APPLICATION NUMBER: 09/918535	?	PRIOR FILING DATE: 1998-04-09
PRIOR FILING DATE: 2001-07-30	?	PRIOR APPLICATION NUMBER: 60/081229
PRIOR APPLICATION NUMBER: 60/062250	?	PRIOR FILING DATE: 1998-04-09
PRIOR FILING DATE: 1997-10-17	?	PRIOR APPLICATION NUMBER: 60/081955
PRIOR APPLICATION NUMBER: 60/064249	?	PRIOR FILING DATE: 1998-04-15
PRIOR FILING DATE: 1997-11-03	?	PRIOR APPLICATION NUMBER: 60/081817
PRIOR APPLICATION NUMBER: 60/065311	?	PRIOR FILING DATE: 1998-04-15
PRIOR FILING DATE: 1997-11-13	?	PRIOR APPLICATION NUMBER: 60/081819
PRIOR APPLICATION NUMBER: 60/066364	?	PRIOR FILING DATE: 1998-04-15
PRIOR FILING DATE: 1997-11-21	?	PRIOR APPLICATION NUMBER: 60/081952
PRIOR APPLICATION NUMBER: 60/077450	?	PRIOR FILING DATE: 1998-04-15
PRIOR FILING DATE: 1998-03-10	?	PRIOR APPLICATION NUMBER: 60/081838
PRIOR APPLICATION NUMBER: 60/077632	?	PRIOR FILING DATE: 1998-04-15
PRIOR FILING DATE: 1998-03-11	?	PRIOR APPLICATION NUMBER: 60/082568
PRIOR APPLICATION NUMBER: 60/077641	?	PRIOR FILING DATE: 1998-04-21
PRIOR FILING DATE: 1998-03-11	?	PRIOR APPLICATION NUMBER: 60/082569
PRIOR APPLICATION NUMBER: 60/077649	?	PRIOR FILING DATE: 1998-04-21
PRIOR FILING DATE: 1998-03-11	?	PRIOR APPLICATION NUMBER: 60/082704
PRIOR APPLICATION NUMBER: 60/077791	?	PRIOR FILING DATE: 1998-04-22
PRIOR FILING DATE: 1998-03-12	?	PRIOR APPLICATION NUMBER: 60/082804
PRIOR APPLICATION NUMBER: 60/078004	?	PRIOR FILING DATE: 1998-04-22
PRIOR FILING DATE: 1998-03-13	?	PRIOR APPLICATION NUMBER: 60/082700
PRIOR APPLICATION NUMBER: 60/078886	?	PRIOR FILING DATE: 1998-04-22
PRIOR FILING DATE: 1998-03-20	?	PRIOR APPLICATION NUMBER: 60/082797
PRIOR APPLICATION NUMBER: 60/078936	?	PRIOR FILING DATE: 1998-04-22
PRIOR FILING DATE: 1998-03-20	?	PRIOR APPLICATION NUMBER: 60/082796
PRIOR APPLICATION NUMBER: 60/078910	?	PRIOR FILING DATE: 1998-04-23
PRIOR FILING DATE: 1998-03-20	?	PRIOR APPLICATION NUMBER: 60/083336
PRIOR APPLICATION NUMBER: 60/078939	?	PRIOR FILING DATE: 1998-04-27
PRIOR FILING DATE: 1998-03-20	?	PRIOR APPLICATION NUMBER: 60/083322
PRIOR APPLICATION NUMBER: 60/079294	?	PRIOR FILING DATE: 1998-04-28
PRIOR FILING DATE: 1998-03-25	?	PRIOR APPLICATION NUMBER: 60/083392
PRIOR APPLICATION NUMBER: 60/079656	?	PRIOR FILING DATE: 1998-04-29
PRIOR FILING DATE: 1998-03-26	?	PRIOR APPLICATION NUMBER: 60/083495
PRIOR APPLICATION NUMBER: 60/079664	?	PRIOR FILING DATE: 1998-04-29
PRIOR FILING DATE: 1998-03-27	?	PRIOR APPLICATION NUMBER: 60/083496
PRIOR APPLICATION NUMBER: 60/079689	?	PRIOR FILING DATE: 1998-04-29
PRIOR FILING DATE: 1998-03-27	?	PRIOR APPLICATION NUMBER: 60/083499
PRIOR APPLICATION NUMBER: 60/079663	?	PRIOR FILING DATE: 1998-04-29
PRIOR FILING DATE: 1998-03-27	?	PRIOR APPLICATION NUMBER: 60/083545
PRIOR APPLICATION NUMBER: 60/079728	?	PRIOR FILING DATE: 1998-04-29
PRIOR FILING DATE: 1998-03-27	?	PRIOR APPLICATION NUMBER: 60/083554
PRIOR APPLICATION NUMBER: 60/079786	?	PRIOR FILING DATE: 1998-04-29
PRIOR FILING DATE: 1998-03-27	?	PRIOR APPLICATION NUMBER: 60/083558
PRIOR APPLICATION NUMBER: 60/079920	?	PRIOR FILING DATE: 1998-04-29
PRIOR FILING DATE: 1998-03-30	?	PRIOR APPLICATION NUMBER: 60/083559
PRIOR APPLICATION NUMBER: 60/079923	?	PRIOR FILING DATE: 1998-04-29
PRIOR FILING DATE: 1998-03-30	?	PRIOR APPLICATION NUMBER: 60/083500
PRIOR APPLICATION NUMBER: 60/080105	?	PRIOR FILING DATE: 1998-04-29
PRIOR FILING DATE: 1998-03-31	?	PRIOR APPLICATION NUMBER: 60/083742
PRIOR APPLICATION NUMBER: 60/080107	?	PRIOR FILING DATE: 1998-04-30
PRIOR FILING DATE: 1998-03-31	?	PRIOR APPLICATION NUMBER: 60/084366
PRIOR APPLICATION NUMBER: 60/080165	?	PRIOR FILING DATE: 1998-05-05
PRIOR FILING DATE: 1998-03-31	?	PRIOR APPLICATION NUMBER: 60/084414
PRIOR APPLICATION NUMBER: 60/080194	?	PRIOR FILING DATE: 1998-05-06
PRIOR FILING DATE: 1998-03-31	?	PRIOR APPLICATION NUMBER: 60/084441
PRIOR APPLICATION NUMBER: 60/080327	?	PRIOR FILING DATE: 1998-05-06
PRIOR FILING DATE: 1998-04-01	?	PRIOR APPLICATION NUMBER: 60/084637
PRIOR APPLICATION NUMBER: 60/080328	?	PRIOR FILING DATE: 1998-05-07
PRIOR FILING DATE: 1998-04-01	?	PRIOR APPLICATION NUMBER: 60/084639
PRIOR APPLICATION NUMBER: 60/080333	?	PRIOR FILING DATE: 1998-05-07
PRIOR FILING DATE: 1998-04-01	?	PRIOR APPLICATION NUMBER: 60/084640
PRIOR APPLICATION NUMBER: 60/080334	?	

PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085689
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697

Alignment Scores:

Aligned. NO.: 1.34e-188 Length: 518
Score: 2687.00 Matches: 518
Percent Similarity: 100.00% Conservativity: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 80.47% Indels: 0
DB: 10 Gaps: 0

US-09-806-194-1 (1-1804) x US-09-978-295A-196 (1-518)

1 ATGGCGCAGCTGGCCCGGGCTGCTGCTGCTGCGCCAGTGGCTCCCGCGGAC 60
|||||
1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20
61 GCCCGGAGCTGGCCCGCGGCTTACGCTGCGCCCTGCGGGTGGCGGCGGCGGCGGAGC 120
|||||
21 AlaProGluLeuAlaProAlaProPheTrpLeuProLeuArgValaAlaAlaAlaTrpAsn 40
121 CGCGTAGTGGCCCGCCCGGGACCCGGGACCCGCGGACCCCTCCGAGCGCCGCGGCGGTG 180
|||||
41 ArgValValAlaProTrpProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60
181 GCGCTGCGCTGGAGCTGCGCTGGCGTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 240
|||||
61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80
241 GTAGAACACCTGCGAGGGGACTCTGGCGGCGGCTACTACCTGGAGATGCTGA-CCGGGACC 300
|||||
81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuLeuGlyTrp 100
301 CCCCCGAGAGCTACAGATTCGTTGACACTGGAGGAGGAGTAACTTTCGCGTGGCGAGGA 360
|||||
101 ProProGlnLysLeuGlnLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120
361 ACCCGGCACTCTACATAGACACGCTACTTTGACACAGAGAGGCTAGCACATACCGGCTCC 420
|||||
121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140
421 AAGGGCTTTCAGCTCAGCTGAAGTACACAGGAGGAGCTGGAGGGCTTCGTTGGGGAA 480
|||||
141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160
481 GACCTGCTACCATCCCAAGAGGCTTCAATACTTCTTTCTTGTCACACAATGCCACTATT 540

161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180
|||||
541 TTTGAATCAGAAATTTCTTTTTCCTGGGATTAATGAATGGAATACTTGGCCTAGCT 600
|||||
181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200
|||||
601 TATGCCACACTTGGCCAAAGCCATCAAGTCTCTGGAGACCTTCTTGGACTCCCTGGTGACA 660
|||||
201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220
|||||
661 CAAGCAACATCCCAACCGTTTCTCCATGCGAGATGTGTGGAGCCGGCTTGGCCGTGCT 720
|||||
221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240
|||||
721 GGATCTGGGACCAACGAGGTACTCTTCTGGTGGGAATTGAACCAAGTTGTGTATAA 780
|||||
241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyIleGluProSerLeuTrpLys 260
|||||
781 GGAGACATCTGCTATACCCCTATTAAAGGAAGAGTGTACTACACAGATAGAAATTCGAAA 840
|||||
261 GlyAspIleTrpTrpThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280
|||||
841 TTGGAAATGGAGGCCAAAGCCTTAATCTGGACTGCGAGAGTATAACCCACAGAGCC 900
|||||
281 LeuGluIleGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300
|||||
901 ATCGTGAGCAGTGGCACACCGCTGCTGCGCCGCGGAGAGTGTTCATCGGTGGTG 960
|||||
301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValAl 320
|||||
961 GAAGCTGTGGCGCGCATCTCTGATCCAGAAATCTCTCATGGTTCTGGACTGGGTCC 1020
|||||
321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340
|||||
1021 CAGCTGGCTGGAGCAATTCGGAAACACCTTGTCTTACTTCCCTAAATCTCCATC 1080
|||||
341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360
|||||
1081 TACCTCAGAGATCAGAACTCCAGCAGCTCATTCGCTATCACAAATCCTGCTCAGCTTAC 1140
|||||
361 TyrLeuArgAspLysAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380
|||||
1141 ATTCAGCCCATGATGGGCGCGGCTTGAATATGAAATGTACCGATTTCGGCATTCGCCA 1200
|||||
381 IleGlnProMetMetGlyAlaGlyLeuAsnIyrGluCysTyrArgPheGlyIleSerPro 400
|||||
1201 TCCACAAATGCGCTGTGATCGTGGTCCGCGGCTGATGGAGGGCTTCTACGTCATCTCGAC 1260
|||||
401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420
|||||
1261 AGAGCCCAAGAGGCTGGGCTTCGCGAGCGAGCCCTGTGCAGAAATTCGAGTGTGTGCA 1320
|||||
421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440
|||||
1321 GTGCTGAAATTCGCGGCTTCTCAACAGAGATGTAGCCAGCAACTGTGTCCCGCT 1380
|||||
441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460
|||||
1381 CAGTCTTTGAGCGAGCCCATTTGTGGAITGTCTCTATCCGCTCATGAGCGTCTGTGGA 1440
|||||
461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480
|||||
1441 GCATCTCTTGTCTTAATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
|||||
481 AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 500
|||||
1501 CGTGACCCCTGAGTCGTCGAATGATGAGTCTCTCTGGTCAGACATCGCTGGA 1554
|||||
501 ArgAspProGluValValAlaAsnAspGluSerSerLeuValArgHisArgTrpLys 518
|||||

; Sequence 2, Application US/09886143					
; Patent No. US2002015999A1					
; GENERAL INFORMATION:					
; APPLICANT: Cordell, Barbara					
; APPLICANT: Schimmoller, Frauke					
; APPLICANT: Liu, Yu-Wang					
; APPLICANT: Quon, Diana Hom					
; TITLE OF INVENTION: Modulation of A Levels by					
; FILE REFERENCE: SCIS.Q22A					
; CURRENT APPLICATION NUMBER: US/09/886.143					
; PRIOR FILING DATE: 2001-06-20					
; PRIOR FILING DATE: 2000-06-28					
; NUMBER OF SEQ ID NOS: 6					
; SOFTWARE: FastSeq for Windows Version 4.0					
; SEQ ID NO 2					
; LENGTH: 518					
; TYPE: PRT					
; ORGANISM: Homo sapiens					
US-09-886-143-2					
Alignment Scores:					
Pred. No.:	1 34e-188	Length:	518		
Score:	2687.00	Matches:	518		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	80.47%	Indels:	0		
DB:	10	Gaps:	0		
US-09-806-194-1 (1-1804) x US-09-886-143-2 (1-518)					
QY	1	ATGGGCGACTGCCCGGCGCTGTGCTCTGTGCGCCAGTGCTCTCTGGCGCC	60		
DB	1	MetGlyAlaLeuAlaArgAlaLeuLeuProLeuAlaGlnTrpLeuLeuArgAla	20		
QY	61	GCCCCGAGTGCCTCCCGCCCTTCACGCTGCCCTCCGCGTGGCGCCGACGAAC	120		
DB	21	AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn	40		
QY	121	CGGTAGTTGCGCCACC CGGAGCCCGGACCCCTGCGAGCGCACCGCAGCGCTTG	180		
DB	41	ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu	60		
QY	181	GCCTGCGCTGAGCCTGCCCTGGGTCCCGCGGGGGCGCGCCAACTTCTTGGCATG	240		
DB	61	AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPhcLeuAlaMet	80		
QY	241	GTAGAACACTCGAGGGGACTCTGCGCGCGGCTACTACTCGAGATGCTGATCGGGACC	300		
DB	81	ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrlrLeuGluMetleuileGlyThr	100		
QY	301	CCCCCGAGAGCTACAGATTCTCGTTGACACTGGAGCAGTAACCTTTGCCGTGCCAGA	360		
DB	101	ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPhcAlaValAlaGly	120		
QY	361	ACCCCGCACTCTACATAGACAGCTACTTGACAGAGAGCTTAGCATATACCGTCC	420		
DB	121	ThrProHisSerTyrlrileaspThrTrpPheAspThrGluArgSerSerThrTyrlrArgSer	140		
QY	421	AAGGGCTTTGACCTCAGCTGAAGTACACACAGGAAGCTGGACGGCTCTCTGGGAAA	480		
Db	201	TyrlaThrLeuAlaLysProSerSerLeuGluThrPhePheAspSerLeuValThr	220		
QY	661	CAAGCAAACATCCCACAACGTTTTTCTCCATGCAGATGTGTGGAGCGCGTTCCTGCT	720		
Db	221	GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla	240		
QY	721	GGATCTGGACCAACGGAGGTAGTCTTGCTTGGTGAATTTGAACCAAGTTGTATAAA	780		
Db	241	GlySerGlyThrAsn:GlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrls	260		
QY	781	GGAGACATCTGCTATACCCCTATTAGGAGAGAGTGTACTACACATAGAAATTCIGAAA	840		
Db	261	GlyAspiletrPrThrProlleLysGluGluTrpTyrlrGlnleGluLeuLys	280		
QY	841	TTCGAAATTTGGAGGCCAAAGCCCTTAATCTGGACTCGACAGAGTATTAAGCCACAGAAGGCC	900		
Db	281	LeuGluileGlyGlnSerLeuAsnLeuAspCysArgGluTyrlrAsnAlaasplysala	300		
QY	901	ATCGTGGACAGTGGCACACCGCTGCTGCCCTGCCCGACAGAGGTGTTGATCGGTGGT	960		
Db	301	IleValaspserGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal	320		
QY	961	GAAGCTGTGCGCGGCATCTCTGATCCAGAAATCTCTGATGGTTCCTGGACTCGSGTCC	1020		
Db	321	GluAlaValAlaArgAlaSerLeulleProGluPheSerAspGlyPheTrpThrGlySer	340		
QY	1021	CAGCTGGCTGTGTSACGAATTCGGAACACCTTGCTTACTTCCCTAAAAATCTCCATC	1080		
Db	341	GlnLeuAlacystprThrAsnSerGluThrProTrpSerTyrlrPheProLysIleSerile	360		
QY	1081	TACTCGAGAGTAGAAGACTCCAGCAGGTTCATTCCTGATACACATCTCCCTCAGCTTAC	1140		
Db	361	TyrLeuArgaspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyrlr	380		
QY	1141	ATTACGCCATGATGGGCGCGCTCGAATTTAGATGTTACCGATTCGGCATTTCCCCA	1200		
Db	381	IleGlnProMetMetGlyAlaGlyLeuAsnTyrlrGlucystTyrlrArgPheGlyIleSerPro	400		
QY	1201	TCCACAAATCGCTGGTGATCGGTGCCCGGTGATGAGGGCTTCTACGTCATCTTCGAC	1260		
Db	401	SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrlrValIlePheAsp	420		
QY	1261	AGAGCCAGAAAGAGGTGGGCTTCGCGAGCGAGCCGCTGTGCAGAAATTCAGCGTCTGCA	1320		
Db	421	ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla	440		
QY	1321	GTGCTGAAATTTCCGGGCGCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCGCT	1380		
Db	441	ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla	460		
QY	1381	CAGTCTTTGAGCAGGCCCATTTTGTGGATGTCTCTATCGCTCATGAGCCTCTGTGGA	1440		
Db	461	GlnSerLeuSerGluProIleLeuTrpIleValSerTyrlrAlaLeuMetSerValCysGly	480		
QY	1441	GCCATCTCTCTGTCTTAATCGCTCTGCTGCTCGCTGCTCGGTGTGAGGTGGCCGC	1500		
Db	481	AlaIleLeuLeuValLeuilleValLeuLeuLeuLeuProPheArgCysGlnArgArgPro	500		
QY	1501	CGTGACCTCAGGTGCTCAATGATGAGTCTCTCTGGTCAGACATCGCTGGGAAA	1554		
Db	501	ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys	518		

[illegible]

APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630PIC27
CURRENT APPLICATION NUMBER: US/09/978,697
CURRENT FILING DATE: 2004-10-16
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077751
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559

QY	241	GTAGACAACCTGCAAGGGGACTCTGGCCGGCGCTACTACTCGAGATGCTGATCGGACC	300
DB	81	ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr	100
QY	301	CCCCCGCAAGCTACAGATTCCTGTTGACACTGGAAGCACTAACTTTGCCGTGGCAGCA	360
DB	101	ProProGlnLysLeuGlnIleLeuValAspThrGlySerAspPheAlaValAlaGly	120
QY	361	ACCCGCGCACTCTACATAGACACTACTTTGACACAGAGAGCTTAGCACATACCGCTCC	420
DB	121	ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer	140
QY	421	AGGGCTTTGACGTCACAGTGAAGTACACACAGGAAGCTGCACGGCTTCGTTGGGGAA	480
DB	141	LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu	160
QY	481	GACCTCGTCACCATCCCAAGGCTCAATACTCTTTCTGTCAACATTGCCACTATT	540
DB	161	AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle	180
QY	541	TTTGAATCAGAGAATTTCTTTTTCCTGGGATTAAATGGAATGGAATCTGGCCCTAGCT	600
DB	181	PheGluSerGluAsnPheLeuProGlyIleGlyTrpAsnGlyIleLeuGlyLeuAla	200
QY	601	TATGCCACACTTGCACGCCATCAAGTTCTCTGGAGACCTCTCTGCATCCCTGCTGACA	660
DB	201	TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr	220
QY	661	CAAGCAACATCCCAAGCTTTTCTCATGCAGATGTGGAGCCGGCTGGCCCTGCTGCT	720
DB	221	GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla	240
QY	721	GGATCTGGGACCAAGGAGGTAGTTGCTCTGGTGGAAATTGAACCAAGTTGTATAAA	780
DB	241	GlySerGlyThrAsnGlyGlySerLeuValLeuGlyIleGluProSerLeuTyrLys	260
QY	781	GGAGACATCTGGTATACCCCTATTAAAGAAAGATGGTACTACAGATAGAAATCTGAAA	840
DB	261	GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTrpGlnIleGluIleLeuLys	280
QY	841	TTGGAAATTGAGGCCAAAGCCTTAATCTGCATGGAGAGATATAACGCACACAAGGCC	900
DB	281	LeuGluIleGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla	300
QY	901	ATCGTGCAGCTGGCACACACCTCTCGCGCTCGCCACAGAAGGTGTTGATCGGTGTG	960
DB	301	IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaVal	320
QY	961	GRAGCTGGCCCGGCATCTCTGATTCACAGATTCCTGTGATGGTTCTGCGACTCGGTC	1020
DB	321	GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer	340
QY	1021	CAGCTGGGGTCTGGACGAATTCGGAACACCTTGGTCTTTCCTTAAATCTCCATC	1080
DB	341	GlnLeuAlaCysTrpIleAsnSerGluThrProTrpSerTyrPheProLysIleSerIle	360
QY	1081	TACCTGAGAGATGAGAACTCCAGCAGGTCTATCCGTATCACAACTCTGCCTCAGCTTAC	1140
DB	361	TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr	380
QY	1141	ATTACGCCCATGATGGGGCGGCTCAANTTATGAATGTTTACCGATTCGGCATTTCCCCA	1200
DB	381	IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro	400
QY	1201	TCCACAATTCGGCTGGTATCGGTGCCACGGTGATGAGGGCTTCTACGTCTATCTTCGAC	1260
DB	401	SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp	420
QY	1261	AGAGCCCAAGAGGGTGGCTTCGCAAGGAGCCCTGTGCAGAAATTCGACGGTGTCTCA	1320
DB	421	ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla	440
QY	1321	GTGTCTGAAATTTCCGGGCGCTTCTTCAACAGAGATGATAGCCAGCAACTGTGTCCCCGCT	1380

```

Alignment Scores:
Pred. No.:      1,348-188
Score:          2637.00
Length:         518
Matches:        518
Percent Similarity: 100.00%
Conservative:   0
Best Local Similarity: 100.00%
Mismatched:     0
Query Match:    80.47%
Indels:         0
DB:             10
Gaps:           0

US-09-806-194-1 (1-1804) x US-09-973-697-196 (1-518)

```


QY 961 GAAGTGTGGCCGGCCGATCTCTGATCCAGAAATCTCTGATGGTTCCTGCAATGGGTC 1020
Db |||||||
QY 321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340
Db |||||||
QY 1021 GAGCTGGCGTGTGACGAAATCGGAACACCTTGCTTACTTCCCTAAATATCCATC 1080
Db |||||||
QY 341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyPheProLysIleSerIle 360
QY 1081 TACCTGAGAGATGAGAACTCCAGCAGGTTCATCCGTATCAACAATCCCTCCAGCTTAC 1140
Db |||||||
QY 361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTy 380
QY 1141 ATTCAGCCCATGATGGGGCGGGCTTCCAGCAGCCCTGATGAATATACCATTCGGCATTCGCCA 1200
Db |||||||
QY 381 IleGlnProMetMetGlyAlaGlyLeuAsnTyArgCysTyArgPheGlyIleSerPro 400
QY 1201 TCCAAATGCGCTGGTGTATCGGTCCAGCGGTGATGGAGGCTCTAGCTGATCTCGAC 1260
Db |||||||
QY 401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGlyProTyTrValIlePheAsp 420
QY 1261 AGAGCCCAAGAGAGGTGGGCTTCCAGCAGCCCTGTCAGAGAAATTCAGGTGGTCCA 1320
Db |||||||
QY 421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGlyIleAlaAla 440
QY 1321 GTGTCGAAATTCGGGCGCTTCTCAACAGAGGATGATGACCAACTGTGTCCCGCT 1380
Db |||||||
QY 441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460
QY 1381 CAGTCTTTGAGGAGCCCATTTGTGGATTGTGCTATGCTGCGCTCATCGAGTCTGTGGA 1440
Db |||||||
QY 461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyraLeuMetSerValCysGly 480
QY 1441 GCCATCTCTTGTCTTAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
Db |||||||
QY 1501 CGTGACCTGAGTCTCAATGATGAGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
Db |||||||
QY 501 ArgAspProGluValValAsnAspGluSerLeuValArgHisArgTrpLys 518
RESULT 12
US-09-999-832A-196
; Sequence 196, Application US/0999832A
; Publication No. US20020192706A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

;	PRIOR APPLICATION NUMBER:	60/084643	
;	PRIOR FILING DATE:	1998-05-07	
;	PRIOR APPLICATION NUMBER:	60/085339	
;	PRIOR FILING DATE:	1998-05-13	
;	PRIOR APPLICATION NUMBER:	60/085338	
;	PRIOR FILING DATE:	1998-05-13	
;	PRIOR APPLICATION NUMBER:	60/085323	
;	PRIOR FILING DATE:	1998-05-13	
;	PRIOR APPLICATION NUMBER:	60/085582	
;	PRIOR FILING DATE:	1998-05-15	
;	PRIOR APPLICATION NUMBER:	60/085700	
;	PRIOR FILING DATE:	1998-05-15	
;	PRIOR APPLICATION NUMBER:	60/085689	
;	PRIOR FILING DATE:	1998-05-15	
;	PRIOR APPLICATION NUMBER:	60/085579	
;	PRIOR FILING DATE:	1998-05-15	
;	PRIOR APPLICATION NUMBER:	60/085580	
;	PRIOR FILING DATE:	1998-05-15	
;	PRIOR APPLICATION NUMBER:	60/085573	
;	PRIOR FILING DATE:	1998-05-15	
;	PRIOR APPLICATION NUMBER:	60/085704	
;	PRIOR FILING DATE:	1998-05-15	
;	PRIOR APPLICATION NUMBER:	60/085697	
Alignment Scores:			
Pred. No.:	1,34e-188	Length:	518
Score:	2687.00	Matches:	518
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	80.47%	Indels:	0
DB:	10	Gaps:	0
US-09-806-194-1 (1-1804) x US-09-999-832A-196 (1-518)			
Qy	1	ATGGGGGCACTGGCCGGGGCGTGGCTGGCTGGCTGGCCAGTGGCTCTCGCGGCC	60
Db	1	MetGlyAlaLeuAlaArgAlaLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla	20
Qy	61	GCCCGGGAGCTGGCCCGCGCCCTTCACGTGCCCTCCGGTGGCGGGCCGACCAAC	120
Db	21	AlaProGluLeuAlaProAlaProPheTrpLeuProLeuArgValAlaAlaThrAsn	40
Qy	121	CGGTAGTTGCCGCCACCCGGGACCCCGGACCCTCGCAGGCCACCGCGAGCGCTTG	180
Db	41	ArgValValAlaProThrProGlyProGlyTrpProAlaGluArgHisAlaAspGlyLeu	60
Qy	181	GGCGTCCCTCGAGCGCTGCCCTGGCGTCCCGCGGGCGCGCAACTCTTGGCCATG	240
Db	61	AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet	80
Qy	241	GTAGACAACCTCAGGGGACTCTGGCCGGGCTACTACTCTGGAGATGCTGATCGGACC	300
Db	81	ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr	100
Qy	301	CCCCGCAGAAGCTACGATTCTGTTCACTCGAGCAGTAACTTTGCCGTGGCAGGA	360
Db	101	ProProGlnGlyLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly	120
Qy	361	ACCCCGCACTCTACATAGACAGCTACTTTTGACACAGAGGTCIAGCACATACCGCTCC	420
Db	121	ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer	140
Qy	421	AAGGGCTTGGAGCTCACAGTGAAGTACACAGAAGCTGGACGGGCTTCGTTGGGGAA	480
Db	141	LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpTrpThrGlyPheValGlyGlu	160
Qy	481	GACCTCGTCACCATCCCCAAGGCTTCAATACTTCTTCTTGTCAACATTTGCCACTATT	540
Db	161	AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle	180
Qy	541	TTTGAATCAGAGAATTTCTTTTGGCTGGGATTAATATGGAATGGAATCTTTGGCCCTAGCT	600
Db	181	PheGluSerGluAsnPhePheLeuProGlyTyleYsTrpAsnGlyIleLeuGlyLeuAla	200

QY 601 TATGCCACATCTGGCAAGCCATCAAGTTCTCTGGAGACCTTCTCGACTCCCTGGTGCACA 669
DB 201 TyrAlaThrLeuAlaLysProSerSerLeuGluThrPhePheAspSerLeuValThr 220
QY 661 CAAGCAACATCCCAAGCTTTCTCCATGCAGATGTGTGAGCGGCTTCCCGGTGCT 720
DB 221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240
QY 721 GGATCTGGGACCAAGGAGGTAGTCTCTCTGGTGGAAATTGAACCAAGTTCCTAAA 780
DB 241 GlySerGlyThrAsnGlyLysSerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260
QY 781 GGACATCTCTGTATACCCCTATTAAAGGAAGAGTGGTACTACCATAGATAAGAAATCTGAAA 840
DB 261 GlyAspIleTyrThrProIleLysGluGluTyrTyrGlnIleGluIleLys 280
QY 841 TTGGAATTTGAGGCCAAGCCTTAATCTGGACTCCAGAGTAGTATACGCAGACAAGGCC 900
DB 281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspCysAla 300
QY 901 ATCGTGGACAGTGGCACCACAGCTGTCTGGCTGCCCCAGAGGTTGTGATGCGGTGGT 960
DB 301 IleValAspSerGlyThrLeuLeuArgLeuProGlnLysValPheAspAlaVal 320
QY 961 GAAGCTGTGGCCCGCATCTCTGATTCAGAAATCTCTGATGGTTTCTGGACTGGCTCC 1020
DB 321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTyrGlySer 340
QY 1021 CAGCTGGGCTGTGGCAAAATTCGGAACACCTGGTGTCTTACTCCCTAAAATCTCCATC 1080
DB 341 GlnLeuAlaCysTyrThrAsnSerGluThrProTyrPheProLysIleSerIle 360
QY 1081 TACCTGAGAGATGAGAACTCCAGAGGTCATCTCGTATCAACATCCCTCCCTCAGCTTAC 1140
DB 361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380
QY 1141 ATTACGCCATGATGGGGCGCGCTGTAATTTATGAATGTTACCGATGGCATTCGCCA 1200
DB 381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400
QY 1201 TCCCAAAATGGCTGTGATCGGTGCCAGGTGATGGAGGCTTCTAGTCATCTCGAC 1260
DB 401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420
QY 1261 AGAGCCCAAGAGAGGTGGCTTCGCCAGCGCCCTGTGCAGAAATTCAGGTGCTGCA 1320
DB 421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluAlaGlyAlaAla 440
QY 1321 GTGTCTGAAATTCGGGCCCTTCTCAGCAGAGATGTAGCCAGCACTGTGTCCCGCT 1380
DB 441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460
QY 1381 CAGTCTTTGAGCGACCCATTTTGTGGATGTGTCTATGCGCTCATGAGGTGTGTGGA 1440
DB 461 GlnSerLeuSerGluProIleLeuThrIleValSerTyrAlaLeuMetSerValCysGly 480
QY 1441 GCCATCTCTGTGTATTATGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1500
DB 481 AlaIleLeuLeuValIleValLeuLeuLeuLeuProPheArgCysGlnArgArgPro 500
QY 1501 CGTACCTGTAGTGTCAATATGATGATGCTCTCTGTGTGCACATCTGTGAAA 1564
DB 501 ArgAspProGluValValAlaAsnAspGluSerSerLeuValArgHisArgIleLys 516

RESULT 13

US-09-978-189-196

; Sequence 196, Application US/09973189

; Publication No. US20030004102A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi

; APPLICANT: Baker, Kevin P.

; APPLICANT: Botstein, David

APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerbier, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James;
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2630P1C7
CURRENT APPLICATION NUMBER: US/09/978,189
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27

/	PRIOR FILING DATE:	1998-04-29
/	PRIOR APPLICATION NUMBER:	60/083559
/	PRIOR FILING DATE:	1998-04-29
/	PRIOR APPLICATION NUMBER:	60/083500
/	PRIOR FILING DATE:	1998-04-29
/	PRIOR APPLICATION NUMBER:	60/083742
/	PRIOR FILING DATE:	1998-04-30
/	PRIOR APPLICATION NUMBER:	60/084366
/	PR-CR FILING DATE:	1998-05-05
/	PRIOR APPLICATION NUMBER:	60/084414
/	PRIOR FILING DATE:	1998-05-06
/	PRIOR APPLICATION NUMBER:	60/084441
/	PRIOR FILING DATE:	1998-05-06
/	PRIOR APPLICATION NUMBER:	60/084637
/	PRIOR FILING DATE:	1998-05-07
/	PRIOR APPLICATION NUMBER:	60/084639
/	PRIOR FILING DATE:	1998-05-07
/	PRIOR APPLICATION NUMBER:	60/084640
/	PRIOR FILING DATE:	1998-05-07
/	PRIOR APPLICATION NUMBER:	60/084598
/	PRIOR FILING DATE:	1998-05-07
/	PRIOR APPLICATION NUMBER:	60/084600
/	PRIOR FILING DATE:	1998-05-07
/	PRIOR APPLICATION NUMBER:	60/084627
/	PRIOR FILING DATE:	1998-05-07
/	PRIOR APPLICATION NUMBER:	60/084643
/	PRIOR FILING DATE:	1998-05-07
/	PRIOR APPLICATION NUMBER:	60/085339
/	PRIOR FILING DATE:	1998-05-13
/	PRIOR APPLICATION NUMBER:	60/085338
/	PRIOR FILING DATE:	1998-05-13
/	PRIOR APPLICATION NUMBER:	60/085323
/	PRIOR FILING DATE:	1998-05-13
/	PRIOR APPLICATION NUMBER:	60/085582
/	PRIOR FILING DATE:	1998-05-15
/	PRIOR APPLICATION NUMBER:	60/085700
/	PRIOR FILING DATE:	1998-05-15
/	PRIOR APPLICATION NUMBER:	60/085689
/	PRIOR FILING DATE:	1998-05-15
/	PRIOR APPLICATION NUMBER:	60/085579
/	PRIOR FILING DATE:	1998-05-15
/	PRIOR APPLICATION NUMBER:	60/085580
/	PRIOR FILING DATE:	1998-05-15
/	PRIOR APPLICATION NUMBER:	60/085573
/	PRIOR FILING DATE:	1998-05-15
/	PRIOR APPLICATION NUMBER:	60/085704
/	PRIOR FILING DATE:	1998-05-15
/	PRIOR APPLICATION NUMBER:	60/085697
Alignment Scores:		
Pred. No.:	1.34e-188	Length: 518
Score:	2687.00	Matches: 518
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	80.47%	Indels: 0
DB:	11	Gaps: 0
US-09-806-194-1 (1-1804) x US-09-978-189-196 (1-518)		
CY	1	ATGGGCGACCTGGCCCGGCGCTGTGCCTCTGTCGCCAGTGGCTCCTTGCGCGC 60
Dd	1	MetGlYAlaLeuAAtArgAlaLeuLeuProLeuAlaGlnTrpLeuLeuArGAla 20
Qy	61	GCCCGGAGCTGGCCCCCGCCCCTTCAcGTGCCCCCTCGGGTGGCGCGGCCAGAAC 120
Dd	21	AlaPrOgluLeuAlaProAlaProPhetrLeuProleuArgValalalaAlatrAsn 40
Qy	121	CSGCTAGTTGGCCACC CGCGGACCCCTCCCGACGCCACGCCGACGCGCTTG 180
Dd	41	ArgValAlaIaprotThrProglyproglyThrProAlaGluArgHisAlaspGlyLeu 60
Qy	181	GGCTCGCCCTGGAGCCTGCCTCGGCTCCCGCGGCGCGCCAActTCttTGGCATG 240

Db 61 AlaleuAlaLeuLupProAlaLeuAlaSerProAlaGlyAlaAlaAsnProLeuAlaMet 60
QY 241 GTACAAACCTGAGGGGGAGCTCTGGCCGGGTACTACCTGAGATGCTGATCGAGACC 300
Db 81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGlnMetLeuIleGlyTyr 190
QY 301 CCCCAGAGAGCTACAGATCTCGTTTCACACCTGGAAGCACTAACTTTGGTGGGAGGA 360
Db 101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnProAlaValAlaGly 120
QY 361 ACCCGACCTCTACATAGACACGTACTTTGACACAGAGAGGCTACACATACCGCTCC 420
Db 121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140
QY 421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCCTCGCTGGGGA 480
Db 141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTyrThrGlyPheValGlyGlu 160
QY 481 GACCTGCTACCATCCCAAGGCTTCAATCTCTTTCTTCTCAACATTCGCCATAT 540
Db 161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 130
QY 541 TTTCGAATCAGAGATTTCTTTTCCCTGGGATTAATGGAAATGGAATCTGGCTAGCT 600
Db 181 PheGluSerGluAsnProPheLeuProGlyIleLysTyrAsnGlyIleLeuGlyLeuAla 200
QY 601 TATGCCACACTTGCCAAAGCATCAAGTTCTCTGAGACCTTCTTCGACTCCCTGGTGACA 660
Db 201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220
QY 661 CAAGCAACATCCCAACGTTTCTCCATGACAGATGTGTGAGCGCGCTTGCCCGTGTCT 720
Db 221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240
QY 721 GGATCTGGGACCAACGAGGTAGTCTTCTTGGTGGTGAATGAACCAAGCTTGTATATA 780
Db 241 GlySerGlyThrAsnGlyLysLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260
QY 781 GGAGACATCTGGTATACCCCTTAATGAAGAGAGTGGTACTACCAATAGAAATCTGAAA 840
Db 261 GlyAspIleTyrThrProLysGluGluTyrTyrGlnIleGluIleLeuLys 280
QY 841 TTGGAATTTGAGGCCAAGCCTTAATCTGGACTGCAGAGATATACGGACAGAGGCC 900
Db 281 LeuGluIleGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300
QY 901 ATCTGTGACAGTGGCACACGCTCTGCGCTGCCAGAGGCTTGTATGGCGTGGIG 960
Db 301 IleValAspSerGlyThrLeuLeuArgLeuProGlnLysValPheAspAlaVal 320
QY 961 GAAGCTGGCCCGCATCTCTGATTCCAGAAATCTCTGAAGTTTCTGGACTGGGTCC 1020
Db 321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTyrThrGlySer 340
QY 1021 CAGCTGGGTCTGGACGAATTCGGAACACCTTGGTCTTACTTCCCTAAATCTCCATC 1080
Db 341 GlnLeuAlaCysTyrThrAsnSerGluThrProTyrPheProLysIleSerIle 360
QY 1081 TACCTGAGAGATGAGAACTCCAGCAGGTCTATCCGTATCAAACTCCCTGACCTTAC 1140
Db 361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380
QY 1141 ATTCAGCCCATGATGGGGCGGCTGAATATGAATGTTTACCGATTGGCATTCGCCA 1200
Db 381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400
QY 1201 TCCACAAATGGCTGGTATCGGTGCCACGGTATGAGAGGCTTCTACGTCACTTCGAC 1260
Db 401 SerThrAsnAlaLeuValIleLeuAlaThrValMetGluGlyPheTyrValIlePheAsp 420
QY 1261 AGAGCCCAAGAGGGTGGGCTTCGACGAGCCCTGTGAGAAATTCAGAGTGGTGA 1320
Db 421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaGala 440

QY 1321 GTCTCTGAATTTCCGGGCCITTTCTCAACAGAGATGTAGCCAGCAACTGTGTCCCGCT 1380
Db 441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460
QY 1381 CAGTCTTTGAGGAGCCCATTTTGTGGATGTGTCTCTATCCGCTCATGAGCGTCTGTGA 1440
Db 461 GlnSerLeuSerGluProIleLeuTyrIleValSerTyrAlaLeuMetSerValCysGly 480
QY 1441 GCATCTCTCTGTCTTAATCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
Db 481 AlaleuLeuValLeuIleValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 500
QY 1501 CGTGACCTCGAGTCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1554
Db 501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518

RESULT 14

US-09-978-608A-196
: Sequence 196, Application US/09978608A
: Publication No. US20030045462A1
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Baker Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleon
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kijavlin, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James;
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: TITLE OF INVENTION: Acids Encoding the Same
: FILE REFERENCE: P2630PLC22
: CURRENT APPLICATION NUMBER: US/09/978,608A
: CURRENT FILING DATE: 2001-10-16
: NUMBER OF SEQ ID NOS: 624
: Prior Application removed - See File Wrapper or Palm
: SEQ ID NO 196
: LENGTH: 518
: TYPE: PRT
: ORGANISM: Homo sapien
US-09-978-608A-196

Alignment Scores:
Pred. No.: 1,34e-188 Length: 518
Score: 2687.00 Matches: 518
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 80.47% Indels: 0
DB: 11 Gaps: 0

US-09-806-194-1 (1-1804) x US-09-978-608A-196 (1-518)

QY 1 ATGGGGCGCACTGGCCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGCC 60

1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20
61 GCGCCGAGTGGCCCGCCGCGCTTACGCTGCCCTCCGGTGGCGCGCCACGAC 120
21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaTrpAsn 40
121 CGGTAGTTCGCGCCCGCCGCGCGCTTACGCTGCCCTCCGGTGGCGCGCCACGAC 180
41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60
181 GCGCTCGCTGGAGCTGCGCTGGCGCTGCCCGCGCGCGCTTACGCTGCCCTCCGGT 240
61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 60
241 GTACACAACTCGAGGGGACTCGCGCGCTTACGCTGCCCTCCGGTGGCGCGCGCTT 300
81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyThrLeuGlnMetLeuLeuGlyThr 200
301 CCGCCGAGAGCTACAGATTCTGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
101 ProGlnLysLeuGlnLeuLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120
361 ACCCGCACTCTACATAGACAGTACTTGCACACAGAGGTCTAGCACATACCGCTCC 420
121 ThrProHisSerThrLeuAspThrThrPheAspThrGluArgSerSerThrThrArgSer 140
421 AAGGCTTTCAGCTCAGTGAAGTACACAAAGGAGTGGACGGCTTCTGGGAA 480
141 LysGlyPheAspValThrValLysThrGlnGlySerThrGlnGlyPheValGlyGln 160
481 GACCTGCTCACCATCCCAAGGCTTCACTTCTTCTGTCACATTCGCTACTT 540
161 AspLeuValThrLeuProLysGlyPheAsnThrSerPheLeuValAsnLeuAlaThrIle 260
541 TTGAATCAGAGAATTTCTTTCCTGGGATTAAATGGAATGGAATTCGCTAGCT 600
181 PheGluSerGluAsnPhePheLeuProGlyLeuLysTrpAsnGlyIleLeuGlyLeuAla 200
601 TATGCCACATTCGCGAGCAACAAGTCTCTGGAGACCTTCTGACATTCCTGGTGA 460
201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220
661 CAACAACTCCCAACGTTTCTCCATGCAGATGTGGAGCGGCTTGGCGCTGCT 720
221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240
721 GGATCTGGACCAACGAGGTAGTCTTGTCTGGTGGAAATTAACCAAGTCTGTATAA 780
241 GlySerGlyThrAsnGlyGlnSerLeuValLeuGlyGlyIleGluProSerLeuThrLys 260
781 GGACATCTGGTATACCTCTATTAGGAAGAGTGGTACTTACCAGATAGAAATCTGA 840
261 GlyAspIleThrThrProLysGluGluThrPheGlnLeuLeuLys 280
841 TTGGAATTTGGAGCCCAAGCTTAAATCTGAGTGCAGAGATATAACGACAGAGGC 900
281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluThrAsnAlaAspLysAla 300
901 ATCTGTCAGATGCGACAGCTGCTGCGCTGCCCGCGCGCTTGTGCTGGTGGT 960
301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaVal 320
961 GAAGCTGTGGCCCGCATCTCTGATTCCAGAAATCTCTGATGGTTCTGCACTGGCTC 1020
321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheThrGlySer 340
1021 CAGTGGCTGCTGGAGCAATTCGGAACACCTTGGCTCTTACTTCCCTAAATCTCCATC 1080
341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerThrPheProLysIleSerIle 360
1081 TACCTGAGATGAGAACTCCAGAGGTCAATCCGCTATCACAATCTGCTGCTGCTTAA 1140

361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380
1141 ATTCAGCCCATGATGGCGCGCTTGAATTAAGATGTAACGATTCGGCATTTCCCA 1200
381 IleGlnProMetMetGlyAlaGlyLeuAsnThrGluCysThrArgPheGlyIleSerPro 400
1201 TCCACAAATTCGCTGATCGGTGCGCGCTTCTCAACAGAGGTAGTACGAGAACTGT 1260
401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheThrValIlePheAsp 420
1261 AGAGCCAGAGAGGTGGCTTCGCGAGGAGCCCTGTGCAGAAATTCAGAGTCTGCA 1320
421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440
1321 GTGCTGCAAAATTCGCGCTTCTCAACAGAGGTAGTACGAGAACTGTCTCCCGCT 1380
441 ValSerGluLeuSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460
1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATGTGCTCTATGCTGCTCATGAGCTCTGGA 1440
461 GlnSerLeuSerGluProIleLeuTrpIleValSerThrAlaLeuMetSerValCysGly 480
1441 GCCATCTCTTCTTAAATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
481 AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuLeuProPheArgCysGlnArgArgPro 500
1501 CGTACCTCTGAGTGTCAATGATGATGCTCTCTCTGCTCAGACATCGCTGGAA 1554
501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518

RESULT 15

US-09-978-585A-196

: Sequence 196, Application US/09978585A

: Publication No. US20030049633A1

: GENERAL INFORMATION:

: APPLICANT: Ashkenazi, Avi

: APPLICANT: Baker Kevin P.

: APPLICANT: Botstein, David

: APPLICANT: Desnoyers, Luc

: APPLICANT: Eaton, Dan

: APPLICANT: Ferrara, Napoleon

: APPLICANT: Filvaroff, Ellen

: APPLICANT: Fong, Sherman

: APPLICANT: Gerber, Hanspeter

: APPLICANT: Goddard, Audrey

: APPLICANT: Godowski, Paul J.

: APPLICANT: Grimaldi, J. Christopher

: APPLICANT: Hillan, Kenneth J.

: APPLICANT: Kijavini, Ivar J.

: APPLICANT: Kuo, Sophia S.

: APPLICANT: Napier, Mary A.

: APPLICANT: Pan, James

: APPLICANT: Paoni, Nicholas F.

: APPLICANT: Roy, Margaret Ann

: APPLICANT: Shelton, David L.

: APPLICANT: Stewart, Timothy A.

: APPLICANT: Tumas, Daniel

: APPLICANT: Williams, P. Mickey

: APPLICANT: Wood, William I.

: TITLE OF INVENTION: Secreted and Transmembrane polypeptides and Nucleic

: FILE OF INVENTION: Acids Encoding the Same

: FILE REFERENCE: P2630P1C15

: CURRENT APPLICATION NUMBER: US/09/978,585A

: CURRENT FILING DATE: 2001-10-16

: NUMBER OF SEQ ID NOS: 624

: Prior Application removed - See File Wrapper or Palm

: SEQ ID NO 196

: LENGTH: 518

: TYPE: PRT

: ORGANISM: Homo sapien


```
QY 181 GCCTCGCCCTGGAGCTCCCTGGGTCCCGGGGGCCGCCCAACTTCTTGGCCATG 240
Db 61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPhelLeuAlaMet 80
QY 241 GTAGACAACCTGCAGGGGACTCTGCCCGGGCTACTACCTGGAGATCCTGATCGGACC 300
Db 81 ValAsnAsnLeuGlnGlyAspSerGlyArgGlyTyrLeuGluMetLeuIleGlyThr 100
QY 301 CCCCGCAGAACTACAGATTCCTGTTGACACTGGAAGCAGTAACCTTCCCTGGCAGGA 360
Db 101 ProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPhelAlaValIleGly 120
QY 361 ACCCGGACTCTACATAGACAGTACTTTGTGACAGAGAGGTCTAGCACATCCGCTCC 420
Db 121 ThrProHisSerTyrIleAspThrIlePheAspThrGluArgSerSerThrTyrArgSer 140
QY 421 AAGGGCTTTGACCTCAGCTGAGTACACACAGGAGCTGGACGGGCTTCCTCGGGA 480
Db 141 LysGlyPheAspValThrValIlyTyrThrGlnGlySerThrIleGlyPheValIleGly 160
QY 481 GACCTCGTCACCATCCCAAGGCTTCAATCTCTTCTGTCAACATTCGCACTAT 540
Db 161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180
QY 541 TTTGAATCAGAGATTTCTTTTGGCTGGGATTAAATGGAATGGAATCTTGGCTAGCT 600
Db 181 PheGluSerGluAsnPhelLeuPheProGlyIleLysIlePheGlyIleLeuGlyLeuAla 200
QY 601 TATGCCACACTTCCCAAGCACTTCTCTCGAGACCTTCTTTCGACTCCCTGGTGACA 660
Db 201 TyrAlaThrLeuAlaLysPrcSerSerLeuGluThrPhePheAspSerLeuValThr 220
QY 661 CAGCAAACTCCCAAGCTTTCTCCATGCAGATGTGGAGCCGGCTGGCGGTGCT 720
Db 221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240
QY 721 GGATCTGGGACCAAGGAGTACTGTCTTGGTGGAAATGAACCAAGTTGTATAA 780
Db 241 GlySerGlyThrAsnGlySerLeuValLeuGlyIleGluProSerLeuTyrLys 260
QY 781 GGACATCTGGTATACCCCTATTAAGGAGAGTGTACTACAGATAGAAATTCGAA 840
Db 261 GlyAspIleTyrThrProLysGlyLeuGluTyrTyrGlnIleGlyIleLeuLys 280
QY 841 TTGGAAATTCGGAGCCCAAGCCTTAATCTGCAGCTGCAGAGTATACGCACACAGGAC 900
Db 281 LeuGluIleGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300
QY 901 ATCGTGACAGTGGCACCGCTGCTGCGCTGCCCGCAGAAAGTGTGTGATCGGTGGG 960
Db 301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320
QY 961 GAAGCTGTGGCCCGGCACTCTGATTCAGAAATTCCTGATGGTITTCGAGCTGGTCC 1020
Db 321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheIleThrGlySer 340
QY 1021 CAGCTGGCGTGTGGAGCAATTCGGAACACCTTGTCTTACTTCCTAAATTCCTCATC 1080
Db 341 GlnLeuAlaCysThrPheAsnSerGluThrProThrSerTyrPheProLysIleSerIle 360
QY 1081 TACTGTAGAGATGAGAACTCCAGCAGGTCAITCCGTATCACAAATCCTGCTCAGCTTAC 1140
Db 361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380
QY 1141 ATTCAGGCCATGATGGGGGCCGCTGAATTAATGTTTACCGATTCGGCAITTCGCCCA 1200
Db 381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400
QY 1201 TCCAAATCGCTGTGTGTCGGTGCACCGGTGATGAGGGCTCTACGTCATCTTCGAC 1260
Db 401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420
QY 1261 AGACCCAGAGAGGGTGGGCTTCGACGAGGCGCCCTGTGTCAGAAATTCAGGCTGTGCA 1320
```

```
Db 421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440
QY 1321 GTCTCTGAAATTTCCGGGCTTTCTCAACAGAGATGTAGCCAGCAACTGTGTCGCCGCT 1380
Db 441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460
QY 1381 CAGTCTTTGAGCGAGCCCAATTTTGTGATTGTGTCTTCATCGCTCATGAGCTCTGTGGA 1440
Db 461 GlnSerLeuSerGluProIleLeuLeuTriPheValSerTyrAlaLeuMetSerValCysGly 480
QY 1441 GCCATCCTCTTGTCTTAATCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
Db 481 AlaIleLeuValLeuIleValLeuLeuLeuLeuProPheArgCysGlnArgArgPro 500
QY 1501 CGTGACCTGAGTGTCAATGATGAGTCTCTCTGCTGACAGATCGTCGAAA 1554
Db 501 ArgAspProGluValValAsnAspGluSerLeuValArgHisArgTrpLys 518

RESULT 3
US-09-548-372D-2
: Sequence 2, Application US/09548372D
: Patent No. 6420534
: GENERAL INFORMATION:
: APPLICANT: GURNEY, ET AL.
: TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND U
: FILE OF INVENTION: THEREOF
: FILE REFERENCE: 29915/62801
: CURRENT APPLICATION NUMBER: US/09/548.372D
: CURRENT FILING DATE: 2000-04-12
: PRIOR APPLICATION NUMBER: US 60/155,491
: PRIOR FILING DATE: 1999-09-23
: PRIOR APPLICATION NUMBER: US 09/404,133
: PRIOR FILING DATE: 1999-09-23
: PRIOR APPLICATION NUMBER: PCT/US99/20881
: PRIOR FILING DATE: 1999-09-23
: PRIOR APPLICATION NUMBER: US 60/101,594
: PRIOR FILING DATE: 1998-09-24
: NUMBER OF SEQ ID NOS: 73
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 2
: LENGTH: 518
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-548-372D-2

Alignment Scores:
Pred. No.: 9,62e-227 Length: 518
Score: 2687.00 Matches: 518
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 80.47% Indels: 0
DB: 4 Gaps: 0

US-09-806-194-1 (1-1804) x US-09-548-372D-2 (1-518)
QY 1 ATGGCGCACTGGCGCGGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
Db 1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 20
QY 61 GCCCGGAGCTGGCGCGCGCGCGCTTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
Db 21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40
QY 121 CGCTAGTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTG 180
Db 41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60
QY 181 GCGCTGCGCGCTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Db 61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPhelLeuAlaMet 80
QY 241 GTAGACAACCTGCAGGGGAGTCTGCGCGCGGCTACTACCTGGAGATGCTGATCGGACC 300
```

```
Db      81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100
Qy      301 CCCCAGAGAGCTACACATCTCCITCACACTGGAACAGTAACTTTGGCGTGGCAGGA 360
Db      101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerAsnPheAlaValAlaGly 120
Qy      361 ACCCGCGACTCCTACATAGACACAGTACCTTGACACAGAGAGTCTAGCACATACGCTCC 420
Db      121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140
Qy      421 AAGGCGTTTGAGCTCACAGTGAAGTACACACAGAGAGCTGGAGCGGCTTCCTGGGGA 480
Db      141 LysGlyPheAspValThrValTysTyrThrGlnGlySerTyrThrGlyPheValGlyGln 160
Qy      481 GACCTCGTCACATCCCAAGAGCTTCATCTCTTCTTGTCACAGATCGCACATATT 540
Db      161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180
Qy      541 TTTGAATCAGAGAAATTTCTTTTGCTCGGATTAATGGAATGGAATCTTGGCTAGCT 600
Db      181 PheGluSerGluAsnPhePheLeuProGlyIleLysTyrAsnGlyIleLeuGlyLeuAla 200
Qy      601 RATGCCACACTTGCACAGCCATCAAGTTCTCTGGAGACTCTCTCGACTCCCTGGTGACA 660
Db      201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220
Qy      661 CAAGCAACATCCCCAAGCTTTCTCCATGCAGATGTCTGGAGCGGCTTCCCGCTTCT 720
Db      221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240
Qy      721 GGATCTGGACCAACGGAGGTAGTCTCTCTGGTGGGAATGAACCAAGTTGTATAAA 780
Db      241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260
Qy      781 GGAGACATCTGTATACCCCTATTAAAGGAAGTGGTACTACCAGATAGAAATCTGAAA 840
Db      261 GlyAspIleTyrThrProIleLysGluGluTyrTyrGlnIleGluIleLeuLys 280
Qy      841 TTGGAATTTGAGGCCAAGCCTTAATCTGGACTGCAGAGTATAACGCACAGAGCC 900
Db      281 LeuGluIleGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300
Qy      901 ATCTGGAGACTGGCACCACGCTGTCTGGCGCTGCCCCAGAGGTGTATGATCGGTGGT 960
Db      301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320
Qy      961 GAAGCTGTGGCCCGCGCATCTCTGATTCAGAAATCTCTGATGGTTCTGGACTGGGTC 1020
Db      321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTyrThrGlySer 340
Qy      1021 CAGCTGGCGTGTGGACCAATTCGGAACACCTTGGTCTTACTTCCCTAAATCTCCATC 1080
Db      341 GlnLeuAlaCysTyrThrAsnSerGluThrProTyrSerTyrPheProLysIleSerIle 360
Qy      1081 TACCTGAGAGATGAGAACTCCAGAGGTATATCCGTATCAACAATCTGCCTAGGTTTAC 1140
Db      361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380
Qy      1141 ATTACAGCCCATGATGGGCGCGCTGGAATATGATCTTACCGATTCCGGCATTTCCCA 1200
Db      381 IleGlnProMetMetGlyValAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400
Qy      1201 TCACAAATGGCTGTGATCGGTGCCCGTGTATGGAGGCTTCTACGTCTATCTTCGAC 1260
Db      401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420
Qy      1261 AGAGCCCAAGAGAGGTGGGCTTCCAGCAGGCCCTCTGCAGAAATTCAGAGTGCACCA 1320
Db      421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440
Qy      1321 GTGTCTGAATTTCCGGGCTTTCTCAACAGAGGATGTAGCCAGCAATGTGTCCCGCT 1380
```

```
1b      441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460
Qy      1381 CAGTCTTTGAGCGAGCCCATTTTGTGATGTGTCTATCGCTATCGCTCATGAGCGTCTGTGA 1440
Db      461 GlnSerLeuSerGluProIleLeuTyrPheIleValSerTyrAlaLeuMetSerValCysGly 480
Qy      1441 GCCATCTCTCTTGTCTTAAATCGTCTCTGCTGCTGCGGTTCGCGGTCTCAGCGTCCGCC 1500
Db      481 AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuProPheArgCysGlnArgArgPro 500
Qy      1501 CGTGACCTCAGGTCTGTCATGATGAGTACTCTCTCTGCTCAGACATCCCTGGAAA 1554
Db      501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTyrLys 518

RESULT 4
US-09-548-367D-2
; Sequence 2, Application US/09548367D
; Patent No. 6446698
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND US
; FILE REFERENCE: 29915/6280H
; CURRENT APPLICATION NUMBER: US/09/548,367D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 518
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-548-367D-2

Alignment Scores:
Pred. NO.:          9,62e-227      Length:          518
Score:             2687.00        Matches:         518
Percent Similarity: 100.00%       Conservative:    0
Best Local Similarity: 100.00%    Mismatches:     0
Query Match:       80.47%         Indels:         0
DB:                4              Gaps:          0

US-09-806-194-1 (1-1804) x US-09-548-367D-2 (1-518)
Qy      1 ATGGCGCACTGGCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGCC 60
Db      1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTyrLeuLeuArgAla 20
Qy      61 GCCCGGAGTGGCCCCCGCGCCCTTCACGCTGCCCTCCCGGTGGCGCGGCCACGAC 120
Db      21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaThrAsn 40
Qy      121 CGCGTAGTTGCGCCCGCCCGGACCCCGGACCCCTCCCGAGCCCGCCGCGCGGCTTG 180
Db      41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60
Qy      181 GCGTCTCCCTCGAGCTGCGCTCGCGGTCCCGCGCGCGCGCGCAACTTCTTTGGCCATG 240
Db      61 AlaLeuAlaLeuGluProAlaLeuAlaSerProIleGlyAlaAlaAsnPheLeuAlaMet 80
Qy      241 GTACACAACCTCGAGGGGAGCTCTGGCGCGGGTACTACTACCTGGAGATGCTGATCGGACC 300
Db      81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100
Qy      301 CCCCAGAGAGCTACAGATCTCTGATCTGACATGGAAGCAGTAGTAACTTTGGCGTGGCAGGA 360
```

```
Db      101  ProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120
QY      361  ACCCGGACCTCTACATAGACAGTACTTTACACAGAGAGGCTAGCACATACCGCTCC 420
Db      121  ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 340
QY      421  AAGGGCTTTGACGTCACAGTGAAGTACACACAAAGGAGCTGGACGGCTTCGTTCGGGAA 480
Db      141  LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 260
QY      481  GACCTGCTCACCATCCCAAGGCTTCAATACTCTTTCTCTCAACATTGGCACTATT 540
Db      161  AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180
QY      541  TTGAATCAGAGAATTCTTTTCCCTGGGATTAAATGGAATGGAATGGAATGGAATGGAAT 600
Db      181  PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200
QY      601  TATGCCACATTTGCCACGACATCAAGTCTCTCGAGACCTCTTCGACTCCCTGGTGACA 660
Db      201  TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220
QY      661  CAAGCAACATCCCAACAGTTTCTCCATGCAGATGTGGACGGCTTGCCCGCTGGCT 720
Db      221  GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240
QY      721  GGATCTGGGACCAACGGAGGTAGCTTGCTGTGGTGGAAATTAACCAAGTTTGTAA 780
Db      241  GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 280
QY      781  GGACACATCTGGTATACCCCTATTAAAGAGAGTGGTACTACAGATAGAAATCTCGAA 340
Db      261  GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrGlnIleGluIleLeuLys 280
QY      841  TTGGAATTTGGAGCCCAAGCTTAATCTGACTGCAGAGATATAACGACAGACAAGCC 900
Db      281  LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300
QY      901  ATCTGGACAGTGGCACACACCTGCTGGGCTGCCCGAGAGGTGTTTGATGGGTG 960
Db      301  IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaVal 320
QY      961  GAAGCTGTGGCCCGCATCTCTGATCCAGAAATCTCTGATGTTTCTGGATGGG 1020
Db      321  GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340
QY      1021  CAGCTGGCGTGTGGACCAATTCGGAACACCTTGCTCTACTTCCCTAAATCTTCATC 1080
Db      341  GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360
QY      1081  TACCTGAGAGATGAGAACTCCAGCAGGTTCATTCGGTATCACAATCTGCTCAGCTTAC 1140
Db      361  TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380
QY      1141  ATTCAGCCCATGATGGGGCGCGCTGAATTATGAATCTTACCGAATTCGGCATTTCCCA 1200
Db      381  IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400
QY      1201  TCCCAAAATGGCTGTGATCGGTGCCACGCTGATGGAGGCTTCTACGTCACTTCGAC 1260
Db      401  SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420
QY      1261  AGAGCCCAAGAGGGTGGCTTCACAGCGAGCCCTGTGCAGAAATTCGAGTGCCTGCA 1320
Db      421  ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440
QY      1321  GTGCTGAATTTCCGGCCCTTTCTCAACAGAGGATGATAGCCCAACTGTGTCGCCCT 1380
Db      441  ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460
QY      1381  CAGTCTTTGAGCGACCCATTTGTGGATTGTGTCCTATGCCCTCATCAGCGCTGTGGA 1440
Db      461  GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480
```

```
QY      1441  GCCATCTCTCTGTCTTAATCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
Db      481  AlalleLeuLeuValLeuIleValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 500
QY      1501  CGTGACCTTGAGTGGTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1554
Db      501  ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518
```

RESULT 5

US-09-551-853D-2

: Sequence 2, Application US/09551853D

: Patent No. 6500667

: GENERAL INFORMATION:

: APPLICANT: GURNEY ET AL.

: TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND L

: FILE REFERENCE: 29915/62801

: CURRENT APPLICATION NUMBER: US/09/551,853D

: PRIOR FILING DATE: 2000-04-18

: PRIOR APPLICATION NUMBER: US 60/155,493

: PRIOR FILING DATE: 1999-09-23

: PRIOR APPLICATION NUMBER: US 09/404,133

: PRIOR FILING DATE: 1999-09-23

: PRIOR APPLICATION NUMBER: PCT/US99/20881

: PRIOR FILING DATE: 1999-09-23

: PRIOR APPLICATION NUMBER: US 60/101,594

: PRIOR FILING DATE: 1998-09-24

: NUMBER OF SEQ ID NOS: 73

: SOFTWARE: PatentIn version 3.1

: SEQ ID NO 2

: LENGTH: 518

: TYPE: PRT

: ORGANISM: Homo sapiens

: US-09-551-853D-2

```
Alignment Scores:
Pred. No.:      9,626-227      Length:      518
Score:          2687.00      Matches:      518
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      80.47%      Indels:      0
DB:               4          Gaps:      0
```

US-09-806-194-2 (1-1804) x US-09-551-853D-2 (1-518)

```
QY      1  ATGGCGGACACTGGCCCGGGGGGCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 60
Db      1  MetGlyAlaLeuAlaArgAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 20
QY      61  GCCCGGAGCTGGCCCGGGCGGCTTACGCTGCGCTCCCGCTCCCGGTTGGCGGCGCCAGAAC 120
Db      21  AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40
QY      121  CCGCTAGTGTGCCCGCCACCGCGGAGCCCGGAGCCCGCTGCGGAGCCCGCGCGCGCGGTG 180
Db      41  ArgValAlaAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60
QY      181  GGGCTCGCCCTGGAGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 240
Db      61  AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80
QY      241  GTAGACAACCTGCAGGGGAGCTCTGCGCGGGCTACTACTCTGGAGATGCTGATCGGAGCC 300
Db      81  ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100
QY      301  CCGCGCGAGAGACCTACAGATTCCTGTCGACTTGGAGACAGTACTACTTTGCGCGCAGGA 360
Db      101  ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120
QY      361  ACCCGGCACTCTACATAGACAGTACTTTCACAGAGAGGCTGTAGCACATACCGCTCC 420
Db      121  ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140
```

```
QY 421 AAGGGCTTIGAGTCACAGTGAAGTACACAAAGCAAGCTGGACGGGCTTCGTGGGAA 480
Db 141 LysGlyPheAspValThrValLysTyrThrGlnGlySerThrGlyPheValGlyGlu 160
QY 481 GACCTGTCACATCCCAAGGCTTCAATACTCTCTTTCTGTGTCACACTGGCACHAT 540
Db 161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180
QY 541 TTGAAACAGACAATTTCTTTTTCCTGGGATTAATCGAATGGANATCTTGGCTACCT 600
Db 181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200
QY 601 TATGCCACACTTGCACCAACCCATCAAGTCTCTGGAGACCTTCTTCGAGCTCCCTGTACA 660
Db 201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220
QY 661 CAAGCAACATCCCAACGCTTTTCTCCATGAGATGTGTGAGCGGGCTTCGCCGTGCT 720
Db 221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240
QY 721 GATCTGGACCAACGGAGGTAGTCTCTTGGTGAATTAAGCAAGTCTGTATAAA 780
Db 241 GlySerGlyThrAsnGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260
QY 781 GGAGACATCTGGTATACCCCTATTAAAGAGAGTGGTACTACCAAGATAGAAATCTGAAA 840
Db 261 GlyAspIleTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280
QY 841 TTGGAATTTGAGGCCAAGGCTTAATCTGGACTGCAGAGATATAAGCGACAGAAGGCC 900
Db 281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300
QY 901 ATCGTGGACAGTGGCACCACCGCTGTGCGCTGCCCAAGAGGTGTTGATCGGCTGGT 960
Db 301 IleValAspSerGlyThrLeuLeuArgLeuProGlnLysValPheAspAlaValAla 320
QY 961 GAAGCTGTGGCCCGCGCATCTCTGATTCAGAAATCTCTGATGGTTTCTGGAGTGGTCC 1020
Db 321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTyrThrGlySer 340
QY 1021 CAGCTGGGCTGTCGACCAATTCGGAACACTTGGTCTTACTTCCCTAAATCTGCATC 1080
Db 341 GlnLeuAlaCysTrpThrAsnSerGluThrProIlePheProLysIleSerIle 360
QY 1081 TACCTGAGAGATGAGAACTCCAGCAGGTTCATTCGATATCAATCTCCCTCAGCTTAC 1140
Db 361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380
QY 1141 ATTCAGCCCATGATGGGGCGGCGCTGAATTAATGATTAATGATTCGGAATTCGCGCA 1200
Db 381 IleGlnProMetMetGlyAlaGlyPheAlaSerProCysTyrArgPheGlyIleSerPro 400
QY 1201 TCCAAATGGCTGTGATCGGTGCCCGGTGTGAGGAGGCTTCTAGTGTATCTTGAC 1260
Db 401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420
QY 1261 AGAGCCCAAGAGAGGTGGGCTTCGACGAGCGCCCTGTGAGAAATTCGAGGTGCTGCA 1320
Db 421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440
QY 1321 GTGCTGAAATTCGGGCGCTTCTTCACACAGAGNTGATGCCACCACTGTGTCGCGCT 1380
Db 441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnGlyValProAla 460
QY 1381 CAGTCTTTGAGGAGCCCATTTTGGGATTCGTGCTATGCGCTCATGAGCGTCTGGGA 1440
Db 461 GlnSerLeuSerGluProIleLeuThrIleValSerTyrAlaLeuMetSerValCysGly 480
QY 1441 GCCATCTCTCTGCTTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
Db 481 AlaIleLeuLeuValIleValLeuLeuLeuLeuProPheArgCysGlnArgArgPro 500
```

```
QY 1501 CGTGACCTGAGGTGCTCAATGATGAGTCTCTGTGTGACAGATCCTCGTGGAA 1554
Db 501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518
RESULT 5
US-09-717-432-2
: Sequence 2, Application US/09717432
: Patent No. 6291223
: GENERAL INFORMATION:
: APPLICANT: ZHU, YUAN
: APPLICANT: LI, XIAOTONG
: APPLICANT: CHRISTIE, GARY
: APPLICANT: POWELL, DAVID J.
: TITLE OF INVENTION: Mouse Aspartic Secretase-1 (mASp1)
: FILE REFERENCE: Gp-70663
: CURRENT APPLICATION NUMBER: US/09/717,432
: CURRENT FILING DATE: 2000-11-21
: PRIOR APPLICATION NUMBER: 60/166,974
: PRIOR FILING DATE: 1999-11-23
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 2
: LENGTH: 514
: TYPE: PKT
: ORGANISM: MUS MUSCULUS
US-09-717-432-2
Alignment Scores:
Pred. No.: 3,43e-201 Length: 514
Score: 2395.00 Matches: 459
Percent Similarity: 92.47% Conservative: 20
Best Local Similarity: 88.61% Mismatches: 35
Query Match: 71.73% Indels: 4
DB: 3 Gaps: 1
US-09-806-194-1 (1-1804) x US-09-717-432-2 (1-514)
```

```
QY 1 ATGGCGCACTGGCGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGCC 60
Db 1 MetGlyAlaLeuLeuArgAlaLeuLeuValLeuAlaGlnTrpLeuLeuSerAla 20
QY 61 GCCCGGAGCTGGCGGGCGGCTTCCAGCTGCCCTCCGGGTGGCGGGCCGACCAAC 120
Db 21 ValProAlaLeuAlaProAlaProPheThrLeuProLeuGlnValAlaGlyAlaThrAsn 40
QY 121 CGGTAGTTGCGCCACCGCGGGACCGCGGACCGCTCCGAGCGCCACGCGCGCTG 180
Db 41 HisArgAlaSerAlaValProGlyLeuGlyThrProGluLeuProArgAlaAspGlyLeu 60
QY 181 GCGTCCGCTGGAGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 240
Db 61 AlaLeuAlaLeuGluProValArgAlaThr-----AlaAsnPheLeuAlaMet 76
QY 241 GTACACAACCTGACGGGGAGCTTGGCGCGGCTACTACCTGGAGATGCTGTGCGGACC 300
Db 77 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 96
QY 301 CCCCCGAGAGCTACAGATTCTCGTTGACACTGGAAGCAGTAGTAACITTCGCTGGCAGGA 360
Db 97 ProProGlnLysValGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 116
QY 361 ACCCGCACTCTACATAGACAGTACTTTGACACAGAGAGGTCTAGCATATCCCGCTCC 420
Db 117 AlaProHisSerTyrIleAspThrTyrPheAspSerGluSerSerSerThrTyrHisSer 136
QY 421 AAGGGCTTTCAGCTCAGTGAAGTACACAAAGCAAGCTGGACGGGCTTCGTGGGAA 480
Db 137 LysGlyPheAspValThrValLysTyrThrGlnGlySerThrGlyPheValGlyGlu 156
QY 481 GACCTGCTGCTGCTTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db 157 AspLeuValThrIleProLysGlyPheAsnSerSerPheLeuValAsnIleAlaThrIle 176
```



```
QY 661 CAAGCAACATCCCAACGTTTCTCCATGCAGATGTCGAGCGCGCTCCCGCTTCT 720
Db 217 GlnAlaLysIleProAspIlePheSerMetGlnMetCysGlyAlaGlyLeuProValAla 236
QY 721 GCATCTGGACCAACGAGGAGTCTTCTGCTGGCTGAAATGAACCAAGTTTGATAAA 750
Db 237 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrIys 256
QY 781 GGAGACATCTGTATACCCCTATTAGGAAGAGTGGTACTACCAAGATAGAAATCTGAAA 840
Db 257 GlyAspIleTyrThrProIleLysGluGluTyrTyrGlnIleGluIleLys 276
QY 841 TTGAAATTTGAGGCCAAGCTTAATCTGCAGCTGCAGAGATATAAGCGACAGAAGGCC 900
Db 277 LeuGluIleGlyGlyGlnAsnLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 296
QY 901 ATCGTGGACATGGCACACCGCTGCTGGCGCTGCCGCCAAGAGTGTTCATCGCGIGGTS 960
Db 297 IleValAspSerGlyThrThrLeuLeuArgIleProGlnLysValPheAspAlaValVal 316
QY 961 GAAGCTGTGGCCCGGCATCTCTGATTCACAAATCTCTGATGGTTCTGCACTGGATCC 1020
Db 317 GluAlaValAlaArgThrSerLeuIleProGluPheSerAspGlyPheTyrThrGlyAla 336
QY 1021 CAGCTGGCGTGTGCAGCAATTCGGAACACCTTGCTTACTTCCCTAAATCTCCATC 1080
Db 337 GlnLeuAlaCysTyrThrAsnSerGluThrProIleAlaTyrPheProLysIleSerIle 356
QY 1081 TACCTGAGAGATGAGAACTCCAGCAGGTCAATTCCTGATATCAAACTCCCTCAGCTTAC 1140
Db 357 TyrLeuArgAspGluAsnAlaSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 376
QY 1141 ATTACGCCCATGATGGGGCGGCTGATTAATGATTAATGATTAATGATTAATGATTAAT 1200
Db 377 IleGlnProMetMetGlyAlaGlyPheAsnTyrGluCysTyrArgPheGlyIleSerSer 396
QY 1201 TCCACAAATGCGTGGTGATCGTGCCACGCTGATGGAGGCTTCTAGCTCATCTTCGAC 1260
Db 397 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValValPheAsp 416
QY 1261 AGAGCCCAAGAGAGGTTGGGCTCCAGCAGGACCCCTGTCGAGAAATTCAGAGTCTGCA 1320
Db 417 ArgAlaGlnArgArgValGlyPheAlaValSerProCysAlaGluIleGluGlyThrThr 436
QY 1321 GTGCTGAAATTTCCGGGCTTCTCAACAGAGGATGAGCAGCAACTGTGTCGCCGCT 1380
Db 437 ValSerGluIleSerGlyProPheSerThrGluAspIleAlaSerAsnCysValProAla 456
QY 1381 CAGCTTTGAGCGAGCCCATTTTGTGGATTGTGTCTATGGCTCATCAGGCTCTGTGGA 1440
Db 457 GlnAlaLeuAsnGluProIleLeuTyrIleValSerTyrAlaLeuMetSerValCysGly 476
QY 1441 GCCATCTCTCTTCTTAATCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
Db 477 AlaIleLeuValLeuIleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 496
QY 1501 CGTGACCTGAGTGGTCAATGATGAGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1554
Db 497 ArgAspProGluValValAsnAspIleSerSerLeuValArgHisArgTyrLys 514
```

RESULT 8

```
US-09-713-158-2
; Sequence 2, Application US/09713158
; Patent No. 6361975
; GENERAL INFORMATION:
; APPLICANT: ZHU, YUAN
; APPLICANT: LI, XIAOTONG
; APPLICANT: POWELL, DAVID J.
; APPLICANT: CHRISTIE, GARY
; TITLE OF INVENTION: MOUSE ASPARTIC SECRETASE-2 (MASP-2)
; FILE REFERENCE: GP-70660
; CURRENT APPLICATION NUMBER: US/09/713,158
```

```
; CURRENT FILING DATE: 2000-11-15
; PRIOR APPLICATION NUMBER: 60/165,800
; PRIOR FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 501
; TYPE: PRT
; ORGANISM: MUS MUSCULUS
; US-09-713-158-2
```

Alignment Scores:

Pred. No.:	1,96e-95	Length:	501
Score:	1186.50	Matches:	238
Percent Similarity:	62.09%	Conservative:	83
Best Local Similarity:	46.03%	Mismatches:	167
Query Match:	35.53%	Indels:	29
DB:	4	Gaps:	8

US-09-806-194-1 (1-1804) x US-09-713-158-2 (1-501)

```
QY 19 GCGTGTGCTGCTGCTCTGCTGGCCAGTGGCTCTCTGCGCGCGCGCGCGAGCTGGCGCCC 78
Db 2 AlaGlnAlaLeuProIlePheLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 18
QY 79 GCGCCCTTCAG-----CTGCCCTCCGGGTGGCGCGCGCGCGCGCGCGCGCGCGCG 123
Db 19 AlaGlnGlyThrHisLeuGlyIleArgLeuProLeuArgSerGlyLeuAla----- 35
QY 124 GTAGTTGCGGCCACCGCGGACCGCCCTGCCAGCGCCACCGCGCGCGCGCGCGCGCG 183
Db 36 -----GlyProLeuGlyLeuArgLeuProArgIleThrAspGluGluSer----- 51
QY 184 CTCGCCCTGGAGCCTGCGCTGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTA 243
Db 52 -----GluGluProGlyArgArgGlySerPheValGluMetVal 64
QY 244 GACAACCTGCGAGGGGACTCTGCGCGCGGCTACTACTGCTGGAGTGTGATCGGCGACCCC 303
Db 65 AsnAsnLeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetIleValGlySerPro 84
QY 304 CCGCAGAGCTACAGATCTCGTTGACACTGGAGAGCAGTAACCTTGGCGTGGCGAGGAAC 363
Db 85 ProGlnIleLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaValGlyAlaAla 104
QY 364 CCGCACTCTACATAGACACGTACTTTGACACAGAGAGGTCTAGACATACCGCTCCAAG 423
Db 105 ProHisProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyrArgAspLeu 124
QY 424 GGCTTTGACCTACAGTGAAGTACACAAAGAGCTGGAGCGGCTTCGTGGGGAAGAC 483
Db 125 ArgLysGlyValTyrValProTyrThrGlnGlyLysTyrGluGlyGluLeuGlyThrAsp 144
QY 484 CTGCTCACCATCCCAAGGCTTCAATCTCTTCTTCTTCTGTCACACATTGCCACTATTTT 543
Db 145 LeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAlaIleThr 164
QY 544 GAATCAGAGAATTTCTTTTTCCTGGGATTAATGAATGGAATGGAATGGAATGGAATG 603
Db 165 GluSerAspLysPhePheIleAsnGlySerAsnTyrGluGlyIleLeuGlyLeuAlaTyr 184
QY 604 GCCACACTTGCACAGCCATCAAGTCTCTCTGAGACCTTCTTCCGACTCCCTGGTGACAAA 663
Db 185 AlaGluIleAlaArgProAspAspSerLeuGluProPhePheAspSerLeuValLysGln 204
QY 664 GCAACATATCCCAAGCTTTCTCCATGCAGATGTGTGGAGCGCGCTTGGCGGTT----- 717
Db 205 ThrHisIleProAsnIlePheSerLeuGlnLeuCysGlyAlaGlyPheProLeuAsnGln 224
QY 718 ---CGTGGATCTGGACCAACGGAGGTAGTCTCTCTTGGTGGTGAATTGAACCAAGTTG 774
Db 225 ThrGluAlaLeuAlaSerValGlyGlySerMetIleIleGlyGlyIleAspHisSerLeu 244
```

QY	775	TATAAGGAGACATCTGGTATACCCCTATTAAAGGAGAGATGGTACTACACAGATAGAAAT	834
Db	245	TyrThrGlySerLeuIrpTyrThrProIleArgArgGluTrpTyrTyrGluValIleIle	254
QY	835	CTGAATTTGGAATTTGGAGGCCAAAGCCCTTAATCTGGACTTCGACAGAGTATAACGCAGAC	894
Db	265	ValArgValGluIleAsnGlyGlnAspLeuLysMetAspCysGlyGluTyrAsnTyrAsp	284
QY	895	AAGCCATCTGTCGACAGTGGCAGCAGCCTCTCGCTGCGCTGCCCAAGAGGTGTTTGALGCG	954
Db	285	LysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysLysValPheGluAla	304
QY	955	GTGGTGGAACTGTGGCCCGCGCATCTCTCATCCAGAATTTCTGTATGGTGTCTCTGGACT	1014
Db	305	AlaValLysSerIleLysAlaAlaSerSerThrGluLysPheProAspGlyPheIrpLeu	324
QY	1015	GGGTCCCGAGCTGGCGTCTGGACGAATTCGGAACACCTTGGCTCTTACTTCCCTAAATC	1074
Db	325	GlyGluGlnLeuValCysTyrGlnAlaGlyThrThrProTrpAsnIlePheProValIle	344
QY	1075	TCCATCTACCTGAGAGATAGAACTCCAGCAGGTCATTCGCGTATCACAATTCCTGCCCTCAG	1134
Db	345	SerLeuTyrLeuMetGlyGlyGluValThrAsnGlnSerPheArgIleThrIleLeuProGln	364
QY	1135	CTTTACATTACGCCCATGATGGGGCGCGCTCGAATTTAT---GAATGTTACCGATTCCGC	1194
Db	365	GlnTyrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTyrLysPheAla	384
QY	1192	ATTTCCTCCATCCCAAAATGGCTGGTGAATCGGTGCCAGCGGTGATGGASGGCTTCTACGTC	1251
Db	385	ValSerGlnSerThrThrGlyThrValMetGlyAlaValIleMetGluGlyPheTyrVal	404
QY	1252	ATCTTCACAGAGCCACAGAGGTGGGTCTCCAGCGGCCCTCTGCAGAAATGCA	1311
Db	405	ValPheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHisValHisAsp	424
QY	1312	GGTGCTCGATGTCTGAAATTTCCGGGCCCTTCTCAACAGAGGATGAGCCAGCAACTGT	1371
Db	425	GluPheArgThrAlaAlaValGluGlyProPheValThrAlaAspMetGluAspCysGly	444
QY	1372	GTCCCCGCTCAGCTTTTCAGCGGCCCATTTTGTGGATTGTGTCTATGCTCATGAGC	1431
Db	445	TyrAsnIleProGlnThrAspGluSerThrLeuMetThrIleAlaTyrValMetAlaAla	464
QY	1432	GTCTGTGGAGCATCTCCTGTGCTTAAATCGCTGCTGGCTGCGCTTCCGGTGTTCAG	1491
Db	465	IleCys---AlaLeuPheMetLeuProLeuCysLeuMetValCysGlnTyrArgCysLeu	483
QY	1492	CGT-----CGCCCGCTGACCTGAGGTCGTCAATCATGACTCTCTCTG	1536
Db	484	ArgCysLeuArgHisGlnHisAspAspPheAlaAspAspIleSerLeuLeu	500

RESULT 9

```

RES001.1
US-09-348-372D-8
; Sequence 8, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/6280I
; CURRENT APPLICATION NUMBER: US/09/548,372D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1

```



```

Db      287  lLevalAspSerGlyThrThrAsnLeuArgLeuProLysLysValPheGluAlaAlaVal 305
QY      961  GAAGCTGTGGCGCGGCATCIGATTCCAGAAATCTCTGAIGCTTCTGGACCTGGGTC 1020
Db      307  LysSerIleGlyAlaAlaSerSerThrGluLysPheProAspGlyPheTrpLeuGlyAla 326
QY      1021  CAGCTGGCTGTGGACGAATTCGGAACACCTTGCTCTTACTTCCCTAAATCTCCATC 1080
Db      327  GluLeuValCysTrpGlnAlaGlyThrThrProTyrPheProValIleSerLeu 346
QY      1081  TACCTGAGAGATGAGAACTCCAGCAGGTATCCGTATCACAAATCTCGCTCAGCTTAC 1140
Db      347  TyrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIleLeuProGlnTyr 366
QY      1141  ATTCAGCCATCATGGGCGCGCTTCTCAAGAGGATTTAT---GAATGTACGATTCCGCAT 1197
Db      367  LeuArgProValGluAspValAlaThrSerGlnAspAspCysTyrLysPheAlaValSer 386
QY      1198  CCATCCACAAATGCGCTGTGTGATCGGTGCCAGGTGATGGAGGCTTCTACCTCATCTC 1257
Db      387  GlnSerSerThrGlyThrValMetGlyAlaValIleMetGluGlyPheTyrValValPhe 406
QY      1258  GACAGAGCCAGAGAGGTTGGCTTCCAGCGAGCCCTGTGCGAAATTCAGAGTGT 1317
Db      407  AspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHisValHisAspGluPhe 426
QY      1318  GCAGTGTCTGAATTTCCGGGCTTCTCAACAGAGGA---GTACCCAGCAGCTGTCTGCC 1377
Db      427  ArgThrAlaAlaValGluGlyProPheValThrAlaAspMetGluAspCysGlyTyrAsn 446
QY      1378  GGTCACTCTTGAGCGACCCATTTTGTGATGTCTCTATCGCTCATGAGCGTCTGT 1437
Db      447  IleProGlnThrAspLysSerThrLeuMetThrIleAlaTyrValMetAlaAlaIleCys 466
QY      1438  GGAGCACTCTCTGTCTATTCGTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1494
Db      467  ---AlaLeuPheMetLeuProLeuLysLeuMetValCysGlnTrpArgCysLeuArgCys 485
QY      1495  ---CGCCCGCGTGCACCGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1536
Db      486  LeuArgHisGlnHisAspAspPheAlaAspAspIleSerLeuLeu 500

RESULT 11
US-09-551-853D-8
; Sequence 8, Application US/09551853D
; Patent No. 6500667
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; FILE OF INVENTION: THEROF
; FILE REFERENCE: 29915/6280L
; CURRENT APPLICATION NUMBER: US/09/551.853D
; CURRENT FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-551-853D-8

Alignment Scores:
Pred. No.: 2,65e-95 Length: 501
Score: 1185.00 Matches: 237
Percent Similarity: 62.14% Conservative: 83

```

```

Best Local Similarity: 46.02% Mismatches: 169
Query Match: 35.49% Indels: 26
DB: 4 Gaps: 7
US-09-806-194-1 (1-1804) x US-09-551-853D-8 (1-501)

QY      25  CTGCTGCCTCTGCTGGCCAGTGGCTCTGCGCGCGCGCGGAGCTGCGCCCGCGGCC 84
Db      1  MetAlaProAlaLeuHisTrpLeuLeuTrpValGlySerGlyMetLeuProAlaGln 20
QY      85  TTACG-----CTGCCCTCGGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 129
Db      21  GlyThrHisLeuGlyIleArgLeuProLeuArgSerGlyLeuAla----- 35
QY      130  GCGCCACCCCGCGGACCCCTCCGAGCGCCACCGCGCGCGCGCGCGCGCGCGCGCGCG 189
Db      36  GlyProLeuGlyLeuArgLeuProArgGluThrAspGluGluser----- 51
QY      190  CTGGAGCCTGCCCTGGCGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 249
Db      52  -----GluGluProGlyArgArgGlySerPheValGluMetValAspAsn 66
QY      250  CTGAGCGCGACTCTGGCGCGCGGTACTACCTGGAGATGCTGATCGGACCGCGCGCGCG 309
Db      57  LeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetThrValGlySerProProGln 86
QY      310  AAGCTACAGATCTCGTTGACACTGGAAGCAGTAACTTTGCGTGGCGAGGACCGCGCAC 369
Db      87  ThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaValGlyAlaAlaProHis 106
QY      370  TCCTACATAGACAGTACTTGTACACAGAGGTCTAGCATATACCGCTCCCAAGGCTTT 429
Db      107  ProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyrArgAspLeuArgLys 126
QY      430  GACGTACAGTGAAGTACACAGAGAGCTGGAGCGCTTCGTTGGGAGAGACCTCGTC 489
Db      127  GlyValTyrProTyrThrGlnGlyLysTrpGluGlyGluLeuGlyThrAspLeuVal 146
QY      490  ACCATCCCAAGGCTTCAATCTCTTTCTGTCACACTTGCACACTATTATTGTAATCA 549
Db      147  SerIleProHisGlyProAsnValThrValArgAlaAsnIleAlaAlaIleThrGluSer 166
QY      550  GAGATTTCTTTTTCCTGGGATTAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 609
Db      167  AspLysPhePheIleAsnGlySerAsnTrpGluGlyIleLeuGlyLeuAlaIleAlaGlu 186
QY      610  CTGCGCAAGCCATCAAGTCTCTGAGACCTTCTTCCACTCCCTGTCGTCACACAAACA 669
Db      187  IleAlaArgProAspAspSerLeuGluProPhePheAspSerLeuValLysGlnThrHis 206
QY      570  ATCCCAACAGTCTTCTCCATGCAGATGTGIGAGCGCGCTTGCCTGTT-----GCT 720
Db      207  IleProAsnIlePheSerLeuGlnLeuCysGlyAlaGlyPheProLeuAsnGlnThrGlu 226
QY      721  GGATCTGGACCAACGAGGTAGTCTTCTTGGTGAATGGAATGGAATGGAATGGAATGGA 780
Db      227  AlaLeuAlaSerValGlyLysSerMetIleIleGlyIleAspHisSerLeuTyrThr 246
QY      781  GGAGCATCTGTATACCCCTATTAAAGGAGAGGTGCTACTACCATAGATAGAAATCTG 840
Db      247  GlySerLeuTyrThrProIleArgArgGluTyrTyrTyrGluValIleIleValArg 266
QY      841  TTGNAATTGGAGCGCAAGCCCTTAATCTGCAGCTGCAGAGAGTATACGCGACGAAGCC 900
Db      267  ValGluIleAsnGlyGlnAspLeuLysMetAspCysLysGluTyrAsnTyrAspLysSer 286
QY      901  ATCGTGGACAGTGGCACCACCTGCTGCGCCTGCCCGAGAGGTGCTTGTATGCGGTGGT 960
Db      287  IleValAspSerGlyThrThrAsnLeuArgLeuProLysValPheGluAlaAlaVal 306
QY      961  GAAGCTGTGGCGCGCGCATCTCTGATCCAGAAATCTCTGATGTTCTGACCTGGGTCC 1020
Db      307  LysSerIleLysAlaAlaSerSerThrGlnLysPheProAspGlyPheTrpLeuGlyGlu 326

```

QY 1021 CAGCTGGCTGCTGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
DB 327 GlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnIlePheProValIleSerIle 346
QY 1081 TACCTGAGAGATGAGAACTCCAGAGGTCAATTCGGTATCATCAATCCCTGGCTGACCTTAC 1140
DB 347 TyrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIleLeuProGlnGlnTrp 366
QY 1141 ATTACGCCCATGATGGGGCCCGGCTCAATAT---GAATGTTACCGATTCGGATTC 1197
DB 367 LeuArgProValIleAspValAlaThrSerGlnAspCysTrpLysPheAlaValSer 1146
QY 1198 CCATCCCAAAATGCGCTGCTGATCGGTGCCACGGTATGAGGAGCTTCTACGTCATCTTC 1257
DB 387 GlnSerThrGlyThrValMetGlyAlaValIleMetGluGlyPheTrpValPhe 406
QY 1258 GACAGAGCCAGAGAGGGTGGCTTCGACGAGGCCCTGTGCGAGAAATTCAGAGTSC 1317
DB 407 AspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHisValHisAspGluPhe 426
QY 1318 GCAGTGTCTGAATTCGCGGCCCTTCTCACACAGAGGATGAGCCCAACTGCTGCC 1377
DB 427 ArgThrAlaAlaValGluGlyProPheValThrAlaAspMetGluAspCysGlyTrpAsn 446
QY 1378 GCTCAGCTTTGAGCGAGCCCAATTTGTGGATTGTGCTCATGCGCTCATGAGGCTGTG 1437
DB 447 IleProGlnThrAspGluSerThrLeuMetThrIleAlaIyrValMetAlaAlaIleCys 466
QY 1438 GGAGCCATCTCTGTGCTTAATCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1494
DB 467 ---AlaLeuPheMetLeuProLeuCysLeuMetValCysGlnTrpArgCysLeuArgCys 485
QY 1495 ---CGCCCCCTGACCTGAGCTGCTGCTCAATCATGATGAGTCTCTCTG 1536
DB 486 LeuArgHisGlnHisAspAspPheAlaAspAspIleSerLeuLeu 500

RESULT 12

US-09-548-372D-4
; Sequence 4, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; FILE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/62801
; CURRENT APPLICATION NUMBER: US/09/548,372D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-548-372D-4

Alignment Scores:

Pred. No.:	9,81e-95	Length:	501
Score:	1178.50	Matches:	240
Percent Similarity:	62.04%	Conservative:	82
Best Local Similarity:	46.24%	Mismatches:	164
Query Match:	35.29%	Indels:	33
DB:	4	Gaps:	9

US-09-806-194-1 (1-1804) x US-09-548-372D-4 (1-501)

QY 19 GCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 78
DB 2 AlaGlnAlaLeuProTrpLeuLeuLeuTrpMet-----GlyAlaGlyValLeuPro 18
QY 79 GCGCCCTTCACG-----CTGCCCTCCGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 123
DB 19 AlaHisClyThrGlnHisGlyIleAlaGlyLeuProLeuArgSerGlyLeuGly----- 35
QY 124 GTAGTGGCGCCCG 183
DB 36 -----GlyAlaProLeu-----GlyLeuArg 42
QY 184 CTGCGCTCGGAG-----CCTGCCCTGGGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 237
DB 43 LeuProArgGluThrAspGluGluProGluGluProGluArgGlySerPheValGlu 62
QY 238 ATGGTACAGAACCTGAGGGGGAGCTCTGGCGCGGGCTACTACCTGGAGATGCTGATCGGG 297
DB 63 MetValAspAsnLeuArgGlyLysSerGlyGlnGlyTyrTrpValGluMetThrValGly 82
QY 298 ACCCGCCCGCAGAGCTACAGNTTCTCGTTGACACTGGAGAGCAGTAACCTTGGCGGTGCA 357
DB 83 SerProProGlnThrLeuAsnIleLeuValAspIleThrGlySerSerAsnPheAlaValGly 102
QY 358 GGAACCCCGCAGCTCTACATAGACAGCTTCTGACACAGAGAGGTCTAGCACATACCGC 417
DB 103 AlaAlaProHisProPheLeuHisArgTyrTrpGlnArgGlnLeuSerSerThrTrpArg 122
QY 418 TCGAAGGGCTTTCAGCTCAGCTGAGTACACACAGAGAGCTGGAGCGGCTTCGTGGG 477
DB 123 AspLeuArgLysGlyValTrpValProTrpThrGlnGlyLysTrpGluGlyGluLeuGly 142
QY 478 GAAGACCTCGCTACCATCCCAAGGCTTCAATACTTCTTCTTCTGTCACATTCGCACT 537
DB 143 ThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAlaAla 162
QY 538 ATTTTGAATCAGAGAAATTTCTTTTTCCTGGGATTAATGGAATGGAATGGAATGGAATG 597
DB 163 IleThrGluSerAspLysPhePheIleAsnGlySerAsnTrpGluGlyIleLeuGlyLeu 182
QY 598 GCTTATGCCACACTTCCCAAGCCATCAAGTCTCTCGAGACCTTCTTCGACTCCCTGGTG 657
DB 183 AlaTrpAlaGluIleAlaArgProAspAspSerLeuGluProPhePheAspSerLeuVal 202
QY 658 ACACAAACAAACATCCCAACGTTTCTCCATGCAGATGTGTGGAGCGGCTGCGCGGT 717
DB 203 LysGlnThrHisValProAsnLeuPheSerLeuGlnLeuCysGlyAlaGlyPheProLeu 222
QY 718 GCTGGATCT-----GGGACCAAGGAGGTAGTCTTCTCTTGGGTGAATTTGAACCA 768
DB 223 AsnGlnSerGluValLeuAlaSerValGlyGlySerMetIleIleGlyGlyIleAspHis 242
QY 769 AGTTGTATAAGGAGACATCTGTATACCCCTATTAAAGGAAGAGGTGTACTACACAGATA 828
DB 243 SerLeuTrpThrGlySerLeuTrpThrProIleArgArgGluTrpTrpTrpTrpTrpTrp 262
QY 829 GAAATCTGAAATTTGGAATTTGGAGCCCAAGGCTTAAATCTGAGTCTGACAGAGTAAAC 888
DB 263 IleIleValArgValGluIleAsnGlyGlnAspLeuLysMetAspCysLysGluTrpAsn 282
QY 889 GCACAGAGCCCATCTGGAGACGTGGCACCACCGCTGCTGCGCTGCCCGCGCGCGCGCGCG 948
DB 283 TyrAspLysSerIleValAspSerGlyThrAsnLeuArgLeuProLysValPhe 302
QY 949 GATCGGCTGGTGAAGCTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1008
DB 303 GluAlaAlaValLysSerIleLysAlaAlaSerSerThrGluLysPheProAspGlyPhe 322
QY 1009 TGGAGTGGGTCGCCAGCTGGCGTGTGGACCAATTCGAAACACCTGGTCTTACTTCCCT 1068
DB 323 TrpLeuGlyGluGlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnIlePhePro 342

Db 363 ProGlnGlnTyrLeuArgProValGluAspValAlaThrSerGluAspSerCysTyrLys 382
QY 1186 TTCGGGATTTCGCCCATCCCAAATGCGTGGTGGATCGGCGACCGGTGATGAGGGCTTC 1245
Db 383 PheAlaIleSerGlnSerThrGlyThrValMetGlyAlaValIleMetGluGlyPhe 402
QY 1246 TAGCTCATCTTCGACAGCCGACAGAGGGTGGCTTCGCGACGAGCCCTGTGCAGAA 1305
Db 403 TyrValValPheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHisVal 422
QY 1306 ATGCAAGGTCCTGCAGTCTGCTGAAATTCCTCCGGGCTTCTCAACAGAGGATGTAGCCAGC 1365
Db 423 HisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGluAsp 442
QY 1366 AACTGTGTCCTCCCGCTCAGCTTGTGACGAGCCCATTTTGTGATGTGTGTCATCGGCTTC 1425
Db 443 CysGlyTyrAsnIleProGlnThrAspGluSerThrLeuMetThrIleAlaTyrValMet 462
QY 1426 ATGAGCGTCTGTGAGCCATCTCTCTTAATCGTCTCTGCTGCTGCTGCTGCTGCTGCTG 1485
Db 463 AlaAlaIleCys---AlaLeuPheMetLeuProLeuCysLeuMetValCysGlnTrpArg 483
QY 1486 TGTCACCGT-----CGCCCCGCTGACCCGAGGTGCTGTAATGATGAGTCCCTCTGCTG 1536
Db 482 CysLeuArgCysLeuArgGlnGlnHisAspAspPheAlaAspIleSerLeuLeuSer 500

RESULT 14

US-09-551-853D-4
; Sequence 4, Application US/09551853D
; Patent No. 6500667
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/6280L
; CURRENT APPLICATION NUMBER: US/09/551,853D
; CURRENT FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-551-853D-4

Alignment Scores:
Pred. No.: 9,8le-95 Length: 501
Score: 1178.50 Matches: 240
Percent Similarity: 62.04% Conservative: 82
Best Local Similarity: 46.24% Mismatches: 164
Query Match: 35.29% Indels: 33
DB: 4 Gaps: 5

US-09-806-194-1 (1-1804) x US-09-551-853D-4 (1-501)

QY 19 GCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 78
Db 2 AlaGlnAlaLeuProTrpLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 18
QY 79 GCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 123
Db 19 AlaHisGlyThrGlnHisGlyIleArgLeuProLeuArgSerGlyLeuGly----- 35
QY 124 GTAGTTGCCGCCACCCCGGAGCCCGGAGCCCTGCGGAGCGCCGACGCCGCGCTTGCGG 183

Db 36 -----GlyAlaProLeu-----GlyLeuArg 42
QY 184 CTCGCCCTGGAG-----CCTGCCCTGGCGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 237
Db 43 LeuProArgGluThrAspGluGluProGluArgGlyArgGlySerPheValGlu 62
QY 238 ATGGTAGACACCTTCACAGGGGACTCTGCGCGCGCTACTACTACCTGCGAGATGCTGATCGGG 297
Db 63 MetValAspAsnLeuArgGlySerGlyGlnGlyTyrTyrValGluMetThrValGly 82
QY 298 ACCCCCCCGCAAGACTACAGATCTCTGCTGACACTGGAAGCAGTAACTTTGCGGIGGCA 357
Db 83 SerProGlnThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaValGly 102
QY 358 GGAACCCCGCTCTCTACATAGACAGCTACTTTGACACAGAGAGTCTAGACATACCGC 417
Db 103 AlaAlaProHisProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyrArg 122
QY 418 TCCAAGGCTTTGAGTGCACAGTGAAGTACACAGGAAGCTGACGGGCTTGCTGGG 477
Db 123 AspLeuArgLysGlyValTyrValProTyrThrGlnGlyLysTrpGluGlyCLeuGly 142
QY 478 GAAGACCTCGTCACCATCCCAAGGCTTCAATCTCTTTCTTGTCAACATTGCCACT 537
Db 143 ThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAlaAla 162
QY 538 ATTTTGTGAATCAGACAAATTTCTTTTGCTGGGATTAATGGAATGGAATCATTGGCCCTA 597
Db 163 IleThrGluSerAspLysPhePheIleAsnGlySerAsnTrpGluGlyIleLeuGlyLeu 182
QY 598 GCTTATGTCACACTGCGCAAGCATCAAGTCTCTGGAGACCTTCTTGGACTCCCTGGT 657
Db 183 AlaIyrAlaGluIleAlaIyrProAspAspSerLeuGluProPhePheAspSerLeuVal 202
QY 658 ACACAAGCAACATCCCAACGCTTTTCTCATGACAGATGTGTGGAGCGCGCTTCCGCTT 717
Db 203 LysGlnThrHisValProAsnLeuPheSerLeuGlnLeuCysGlyAlaGlyPheProLeu 222
QY 718 GCTGGATCT-----GGGACCAACGAGGTAGTCTTGTCTGGGTGGAATGAACCA 768
Db 223 AsnGlnSerGluValLeuAlaSerValGlySerMetIleIleGlyGlyIleAspHis 242
QY 769 AGTTTGTATAAGGAGACATCTGTATACCCCTATTAAAGGAAGAGTGTACTACACATA 828
Db 243 SerLeuTyrThrGlySerLeuTrpTyrThrProIleArgAlaGluTrpTyrTyrGluVal 262
QY 829 GAATTCCTGAATTTGGAAATTTGGAGGCCAAACCTTAATCTGGACTGACAGAGATTAAC 888
Db 263 IleIleValArgValGluIleAsnGlyGlnAspLeuLysMetAspCysLysGluTyrAsn 282
QY 889 GCAGACAAGGCATCGTGACAGTGGCACACGCTGCTGCGCGCTGCCCGCCACAGAGTGT 948
Db 283 TyrAspLysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysLysValPhe 302
QY 949 CATCGCGGTGGTGAAGCTGTGGCGCGCGCTCTCTGATTCAGAAATTCCTGATGGTTT 1008
Db 303 GluAlaAlaValLysSerIleLysAlaAlaSerThrGluLysPheProAspGlyPhe 322
QY 1009 TGGACTGSGTCCGACGCTGCGTGGAGCAATTCGGAACACACCTTGGTCTTACTTCCT 1068
Db 323 TrpLeuGlyGluGlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnIlePhePro 342
QY 1069 AAAATCTCATCTACTCGAGAGATGAGAACTCCAGCAGTCAATCCGCTATCACAACTCTG 1128
Db 343 ValIleSerLeuTyrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIleLeu 362
QY 1129 CCTCAGCTTTACATTCAGCCCATGATGGGGCGCGCTGAAATAT---GAAATGTACCGA 1185
Db 363 ProGlnGlnTyrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTyrLys 382
QY 1186 TTCGCGCATTTCCCATCCACAAATGCGCTGTGTGTCGCGCAGCGGTGATGAGCGGCTTC 1245

```

Db      383 PheAlaIleSerGlnSerSerThrGlyThrValMetGlyAlaValAlloMetGluGlyPhe 402
QY      1246 TAGGTCATCTTCAGACAGCCAGAGAGGTGGGCTTGGCAGCAGAGCCCTGTGCACAA 1375
Db      403 TyrValValPheAspArgAlaArgysArgileGlyPheAlaValSerAlaCysHisVal 422
QY      1306 ATTGCAGGTGCTGCAGTGTCTCAATTTCCGGCCCTTCTCAACAGAGATGACCCAGC 1365
Db      423 HisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGluAsp 442
QY      1366 AACTGTGTCCCGCTCAGCTTGTGAGCAGCCCAATTTCGTGGATTGTGCTATCGCGCTC 1425
Db      443 CysGlyTyrAsnIleProGlnThrAspGluSerThrLeuMetThrIleAlaTyrValMet 462
QY      1426 ATGAGGCTGTGGAGCCATCTCTGTCTTAATCGCTGCTGCTGCTGCTGCTGCTGCTG 1485
Db      463 AlaAlaIleCys---AlaLeuPheMetLeuProLeuCysLeuMetValCysIntrParg 481
QY      1486 TGTACAGCT-----CGCCCGCTGACCTGAGGTGCTGCAATGATGAGTCTCTCTG 1536
Db      482 CysLeuArgCysLeuArgGlnGlnHisAspPheAlaAspIleSerLeuLeu 500

RESULT 15
US-09-604-608-3
; Sequence 3, Application US/09604608
; Patent No. 6545127
; GENERAL INFORMATION:
; APPLICANT: Tang, Jordan J.N.
; APPLICANT: Lin, Xinli
; APPLICANT: Koelsch, Gerald
; TITLE OF INVENTION: Catalytically Active Recombinant Mernapsin and Methods
; TITLE OF INVENTION: of Use Thereof
; FILE REFERENCE: OMRF 179
; CURRENT APPLICATION NUMBER: US/09/604,608
; PRIOR FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: 60/141,363
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: 60/168,060
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: 60/177,836
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 60/178,368
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 60/210,292
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Pro-memapsin 2
; OTHER INFORMATION: Amino Acids 1-15 are vector-derived residues
; OTHER INFORMATION: Amino Acids 16-64 are a putative pro peptide
; OTHER INFORMATION: Amino Acids 1-13 are the T7 promoter
; OTHER INFORMATION: Amino Acids 16-456 are Pro-memapsin 2-T1
; OTHER INFORMATION: Amino Acids 16-421 are Promemapsin 2-T2
US-09-604-608-3

Alignment Scores:
Pred. No.: 1,8e-94 Length: 503
Score: 1175.50 Matches: 235
Percent Similarity: 63.97% Conservative: 81
Best Local Similarity: 47.57% Mismatches: 160
Query Match: 35.21% Indels: 18
DB: 4 Gaps: 6

US-09-806-194-1 (1-1804) x US-09-604-608-3 (1-503)
QY      105 GCGCGGGCCAGACCGGGTAGTGGCGCCACCCGCGGACCCCGGACCCCTGCCGAGC 164
Db      11 GlyArgGlySerMetAlaGlyValLeuProAlaHisGlyThrGlnHisGlyIleArgLeu 30

```

```

QY      165 CCA---CGCCGACGGCTTGGCGCTCGCCCTGGAGCCCTGC----- 200
Db      31 ProLeuArgSerGlyLeuGlyGlyAlaProLeuGlyLeuArgLeuProArgGluThrAsp 50
QY      201 -----CCTGGCGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 254
Db      51 GluGluProGluGluProGlyArgArg-GlySerPheValGluMetValAspAsnLeuAr 70
QY      255 GGGGAGCTCTGGCGCGCGCTACTACCTGGAGATGCTGCGGAGACCCCGCGCGCGCGCGCGCG 314
Db      70 gGlyLysSerGlyGngLysTyrValGluMetThrValGlySerProGlnThrLe 90
QY      315 ACAGATCTCGTTGACACTGGAGCAGTAACCTTTCGCGTGGCAGGAACCCCGCGCACTCCTA 374
Db      90 uAsnIleLeuValAspThrGlySerSerAsnPheAlaValGlyAlaAlaProHisProPh 110
QY      375 CATAGACAGCTACTTTCACACAGAGAGGTCTAGCACATACCGCTCCAAGGCTTTGACGT 434
Db      110 eLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyrArgAspLeuArgLysGlyVa 130
QY      435 CACAGTGAAGTACACACAGGAAGCTGGACGGGCTTCGTTGGGAGACACCTCGCTCACCAT 494
Db      130 lTyrValProTyrThrGlnGlyLysTyrGluGlyGluLeuGlyThrAspLeuValSerIl 150
QY      495 CCCCAAGGCTTCAATACTTCTTTCTTCGACATTCGCCACTATTTTTGAATCAGAGAA 554
Db      150 eProHisGlyProAsnValThrValArgAlaAsnIleAlaAlaIleThrGluSerAspLy 170
QY      555 TTTCTTTTGGCTGGATTAAATGGAATACTTGGCTAGCTTATGTCACACTTCG 614
Db      170 sPhePheIleAsnGlySerAsnTyrGluGlyIleLeuGlyLeuAlaIleAlaGluIleAl 190
QY      615 CAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGIGACACAAACAACATCCC 674
Db      190 aArgProAspAspSerLeuGluProPhePheAspSerLeuValLysGlnThrHisValPr 210
QY      675 CAAGCTTTTCTCCATGCAGATGTGTGGAGCGGCTTCGCCGTTGCTGGATCT----- 726
Db      210 oAsnLeuPheSerLeuGlnLeuCysGlyAlaGlyPheProLeuAsnGlnSerGluValLe 230
QY      727 -GGGACCAACGAGGTAGTCTTGTCTTGGGTGAATGAACCAAGTTGTATAAAGGAGA 785
Db      230 uAlaSerVa-GlyGlySerMetIleIleGlyGlyIleAspHisSerLeuTyrThrGlySe 250
QY      786 CATCTGGTATACCCCTATTAGGAAGAGTGTACTACCATGAGAAATCTGAAATTGGA 845
Db      250 rLeuTrpTyrThrProIleArgArgGluTyrTyrGluValIleIleValArgValGl 270
QY      846 AATTGGAGGCCAAAGCCTTAATCTGACTGCAGAGATATAACCCACAGAGCCATCGT 905
Db      270 uIleAsnGlyGlnAspLeuLysMetAspCysLysGluTyrAsnTyrAspLysSerIleVa 290
QY      906 GGACAGTGGCACCACCGTGTGCGCCCTGCCCGCAGAGAGTGTTCGATCGGTGGTGAAGC 965
Db      290 LaspSerGlyThrThrAsnLeuArgLeuProLysLysValPheGluAlaAlaValLysSe 310
QY      966 TGTGGCGCGCGCATCTCTGATTCAGAAATCTCTGATGCTTCTGGACTGGTCCCGCCAGCT 1025
Db      310 rIleLysAlaAlaSerSerThrGluLysPheProAspGlyPheTrpLeuGlyGluGlnLe 330
QY      1026 GCGCTGTGGAGGAATTCGGAACACACCTTGGCTCTTACTTCCTCCCTAAATCTCCATACCT 1085
Db      330 uAlcYstrPrlGlnAlaGlyThrThrProTyrPasnIlePheProValIleSerLeuTyrLe 350
QY      1086 GAGATGAGAAATCTCCAGCAGGTGCTTCCCGTATCACAAATCTCGCTCAGCTTTACATTCA 1145
Db      350 uMetGlyGluValThrAsnGlnSerPheArgIleThrIleLeuProGlnGlnTyrLeuAr 370
QY      1146 CCCCATGATGGGCGCGCGCTGAAATAT---GAATGTTACCGATTCCGCATTCGCCCATC 1202
Db      370 gProValGluAspValAlaThrSerGlnAspAspCysTyrLysPheAlaIleSerGlnSe 390

```

```

Qy 1203 CACAAATGCGTGTGATCGTCACGAGGTGAGGGGCTTCTAGTCATCTTCGACAG 1262
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
390 rSerThrGlyThrValMetGlyAlaValIleMetGluGlyPheTyrValValPheAspAr 410
Qy 1263 AGCCACAGAGGGTGGGCTTCGACGAGCGCCCTGTGCAGAAATTGCAGGTGCTGCAG 1322
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:
410 gAlaArgLysArgIleGlyPheAlaValSerAlaCysHisValHisAspGluPheArgTh 430
Qy 1323 GTCGTGAATTCGGGGCTTCTCTACAGAGGATGTAGCCAGCACTGTGCCCGCTCA 1382
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:
430 rAlaAlaValGluGlyProPheValThrLeuAspMetGluAspCysGlyTyrAsnIlePr 450
Qy 1383 GTCCTTTGAGCGAGCCCATTTTGTGGATTGTCTCTATCGCTCATGAGCGCTGTGAGG 1442
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:
450 oGlnThrAspGluSerThrLeuMetThrIleAlaTyrValMetAlaAlaIleCys---Al 459
Qy 1443 CATCTCTCTTGTCTTAATCGTCGCTGCTGCTCCGCTCCGGTGTCCAGCGT-----CG 1496
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:
469 aLeuPheMetLeuProLeuCysLeuMetValCysGlnTrpArgCysLeuArgCysLeuAr 489
Qy 1497 CCCCCTGACCCCTGAGGTGCTCATGATGATGCTCTCTG 1536
Db      |||:|||||:|||||:|||||:|||||:|||||:
489 gGlnGlnHisAspPheAlaAspIleSerLeuLeu 502

```

Search completed: September 25, 2003, 11:10:56
Job time : 62.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: September 25, 2003, 10:19:24 ; Search time 114.5 Seconds
(without alignments)
5001.620 Million cell updates/sec

Title: US-09-806-194-1

Perfect score: 339

Sequence: 1 atgggcgactggccgggc.....aaaccccaaaaaaaaaa 1804

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
delop 6.0 , Delext 7.0

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 2215726

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame-n2p.model -DEV=xlp
-Q=/cgn2_1/USPO.spool/US09806194/runat_25092003_095453_10998/app_query.fasta_1.1951
-DB=A_Geneseq_19Jun03 -OFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=5
-MODE=LOCAL -OUTFTW=ptw -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09806194.qcgn_1.158 -brunat_25092003_095453_10998 -NCPU=6 -ICPU=3
-NO_WMAP -LARGESQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_19Jun03.*

1: /SIDSL/gcgdata/gcnseq/gcnseqq-emb1/AA1380.DAT:*
2: /SIDSL/gcgdata/gcnseq/gcnseqq-emb1/AA1381.DAT:*
3: /SIDSL/gcgdata/gcnseq/gcnseqq-emb1/AA1382.DAT:*
4: /SIDSL/gcgdata/gcnseq/gcnseqq-emb1/AA1383.DAT:*
5: /SIDSL/gcgdata/gcnseq/gcnseqq-emb1/AA1384.DAT:*
6: /SIDSL/gcgdata/gcnseq/gcnseqq-emb1/AA1385.DAT:*
7: /SIDSL/gcgdata/gcnseq/gcnseqq-emb1/AA1386.DAT:*
8: /SIDSL/gcgdata/gcnseq/gcnseqq-emb1/AA1387.DAT:*
9: /SIDSL/gcgdata/gcnseq/gcnseqq-emb1/AA1388.DAT:*
10: /SIDSL/gcgdata/gcnseq/gcnseqq-emb1/AA1389.DAT:*
11: /SIDSL/gcgdata/gcnseq/gcnseqq-emb1/AA1390.DAT:*
12: /SIDSL/gcgdata/gcnseq/gcnseqq-emb1/AA1391.DAT:*
13: /SIDSL/gcgdata/gcnseq/gcnseqq-emb1/AA1392.DAT:*
14: /SIDSL/gcgdata/gcnseq/gcnseqq-emb1/AA1393.DAT:*
15: /SIDSL/gcgdata/gcnseq/gcnseqq-emb1/AA1394.DAT:*
16: /SIDSL/gcgdata/gcnseq/gcnseqq-emb1/AA1395.DAT:*
17: /SIDSL/gcgdata/gcnseq/gcnseqq-emb1/AA1396.DAT:*
18: /SIDSL/gcgdata/gcnseq/gcnseqq-emb1/AA1397.DAT:*
19: /SIDSL/gcgdata/gcnseq/gcnseqq-emb1/AA1398.DAT:*
20: /SIDSL/gcgdata/gcnseq/gcnseqq-emb1/AA1399.DAT:*
21: /SIDSL/gcgdata/gcnseq/gcnseqq-emb1/AA1400.DAT:*
22: /SIDSL/gcgdata/gcnseq/gcnseqq-emb1/AA2001.DAT:*
23: /SIDSL/gcgdata/gcnseq/gcnseqq-emb1/AA2002.DAT:*
24: /SIDSL/gcgdata/gcnseq/gcnseqq-emb1/AA2003.DAT:*

Printed. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2687	80.5	518	19	Aspartic proteinas
2	2687	80.5	518	20	Human PRO852 prote
3	2687	80.5	518	20	Human CSP56, aspar
4	2687	80.5	518	20	Human aspartyl pro
5	2687	80.5	518	21	Human PRO852 (UN04
6	2687	80.5	518	21	Human aspartyl pro
7	2687	80.5	518	22	Human-Asp 1 protei
8	2687	80.5	518	22	Human-Asp 1 protei
9	2687	80.5	518	22	Human PRO polypt
10	2687	80.5	518	22	Human aspartyl pro
11	2687	80.5	518	22	Human aspartyl pro
12	2687	80.5	518	22	Human aspartyl pro
13	2687	80.5	518	22	Human aspartyl pro
14	2687	80.5	518	22	Human Asp-1 protei
15	2687	80.5	518	23	Human Asp-1 protei
16	2687	80.5	518	23	Human Asp-1 protei
17	2687	80.5	518	23	Human aspartyl pro
18	2687	80.5	518	23	Human BACE2 amino
19	2687	80.5	518	24	Human PRO852 prote
20	2687	80.5	518	24	Human secreted/tira
21	2687	80.5	518	24	Novel human secret
22	2687	80.5	518	24	Human secreted/tira
23	2687	80.5	518	24	Human PRO852 poly
24	2687	80.5	518	24	Human PRO polypt
25	2687	80.5	518	24	Human PRO polypt
26	2687	80.5	518	24	Human secreted/tira
27	2687	80.5	518	24	Human PRO polypt
28	2687	80.5	518	24	Human secreted/tira
29	2506	75.1	481	22	Human colon cancer
30	2437	73.0	475	22	Secreted recombin
31	2437	73.0	475	22	Human secreted asp
32	2437	73.0	475	22	Secreted recombin
33	2395	71.7	514	22	Amino acid sequenc
34	2293	68.7	439	23	Human polypeptide
35	2207	66.1	423	22	Human aspartyl pro
36	2184	65.4	423	22	Human membrane or
37	2124	63.6	413	22	Acid-processed hu-
38	2124	63.6	413	22	Human acid-process
39	2124	63.6	413	23	Asp-1deltaTM(his)6
40	1794	53.7	355	22	Human protein sequ
41	1187	35.5	501	21	Rat beta-secretase
42	1186.5	35.5	501	22	Mouse aspartic sec
43	1185	35.5	501	21	Murine beta-secret
44	1185	35.5	501	21	Murine aspartyl pr
45	1185	35.5	501	22	Murine aspartyl pr

ALIGNMENTS

RESULT 1
AAW61362
ID AAW61362 standard; Protein; 518 AA.
XX
AC AAW61362;
XX
DT 25-MAR-2003 (updated)
DT 25-SEP-1998 (first entry)
DE
DE Aspartic proteinase ASPI.
XX
KW ASPI; Aspartic proteinase; Alzheimer's disease; cancer; melanoma.
XX
OS Homo sapiens.
XX
PN EP648062-A2.
XX
PD 17-JUN-1998.

```

XX
XX PF 01-DEC-1997; 97EP-0309648.
XX
XX PR 14-DEC-1996; 96GB-0036022.
XX PR 06-OCT-1997; 97US-0999723.
XX
XX PA (SMIK ) SMITHKLINE BEECHAM CORP.
XX PA (SMIK ) SMITHKLINE BEECHAM PLC.
XX
XX PI Chapman CG, Evans JR, Powell DJ, Southan C:
XX
XX WPI; 1998-314477/28.
XX DR N-PSDB; AAV27962.
XX
XX PT New isolated polynucleotide encodes Aspartic protease polypeptide -
XX PT used to diagnosis, treat and vaccinate against Alzheimer's disease,
XX PT cancer and melanoma
XX
XX PS Claim 11; Page 7; 19pp; English.
XX
XX CC The human ASP1 protein is structurally related to other proteins of the
XX CC Aspartic proteinase family. ASP 1 polypeptides and polynucleotides can
XX CC be used to diagnosis, treat and vaccinate against Alzheimer's disease,
XX CC cancer and melanoma
XX CC (Updated on 25-MAR-2003 to correct PR field.)
XX
XX SQ Sequence 518 AA:
XX
XX Alignment Scores:
XX Pred. No.: 4,01e-213 Length: 518
XX Score: 2687.00 Matches: 518
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 80.47% Indels: 0
XX DB: 19 Gaps: 0
XX
XX US-09-806-194-1 (1-1804) x AAW61362 (1-518)

```


Db 1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20
Qy 61 GCCCGGAGCTGGCCCGCCGCTTCACGCTCCGCTCGCGTGGCCGCGCCACGAC 120
Db 21 AlaProGluLeuAlaProAlaProPheThrLeuProGluArgValAlaAlaIleThrAsn 40
Qy 121 CGGTAGTTCGGCCACCGCCGACCGCGGACCCCTGGCAGCGCCACCGCCAGCGCTTG 180
Db 41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60
Qy 181 GCGCTCCCTGGAGCTGCCCTGGCTCCCGCGCGCGCGCCCAACTCTTGGCATG 240
Db 61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80
Qy 241 GTAGACAACCTGAGGGGACTCTGGCGGGGCTACTACCTGAGATGCTGATCGGAC 300
Db 81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100
Qy 301 CCGCGGAGAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACCTTCCCTGGCAGCA 360
Db 101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120
Qy 361 ACCCGCACTCTACATAGACAGCTACTTTGACAGAGAGGTCTAGCACATCCGCTCC 420
Db 121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140
Qy 421 AAGGGCTTTCAGCTCAGTCAAGTACACACAGGAGGCTGGACGGCTTCGTTGGGAA 480
Db 141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160
Qy 481 GACTCTGCTACCATCCCAAGCCTCAATACTTCTTGTCAACATTCGCACTAT 540
Db 161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180
Qy 541 TTGAATCAGAGAATTCTTTTGGCTGGATTAATGGAAATGGAAATCTGSCCTACT 600
Db 181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200
Qy 601 TATGCCACTTCCCAAGCCACTCAAGTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
Db 201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220
Qy 661 CAAGCAACATCCCAACGTTTCTTCATCCAGCATGTGGAGCGCGCTGGCGTGGT 720
Db 221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240
Qy 721 GGATCTGGGACCAACGAGGTAGTCTTGTCTGGTGGAAATGAACCAAGTTTGTATAA 780
Db 241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260
Qy 781 GGACACATCTGGTATACCCCTNTAGGAAGAGTGGTACTACAGATAGAAATCTGAMA 840
Db 261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluLeuLys 280
Qy 841 TTGGAATTTGAGGCCAAAGCCTTAATCTGCAGTGCAGAGATATAACGACAGAGGCC 900
Db 281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300
Qy 901 ATCTGGACAGTGCACACCGCTGCTCCGCGCTGCCCGCAGAGGTGTTCATCGGTGGT 960
Db 301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaVal 320
Qy 961 GAAGCTGGCCCGCATCTCTGATTCAGATTCCTCATGATTCCTGGACTGGGTCC 1020
Db 321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlyLys 340
Qy 1021 CAGCTGGCTGCTGGACCAATTCGGAACACCTTGTCTTACTTCCTCAAAATCTCCATC 1080
Db 341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360
Qy 1081 TACTGTAGATGAGAACTCCACAGGTCAATCCGTATCACATCTCGCTCAGCTTAC 1140
Db 361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy 1141 ATTCAAGCCATGATGGGCGCGCTCAATTGAATTAATGATTCGGCATTCGCCCA 1200
Db 381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCystTyrArgPheGlyIleSerPro 400
Qy 1201 TCCACAAATCGGCTGGTATCGGTGGCCACGGTATGAGAGGGCTTCTACGTCACTTCGAC 1260
Db 401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420
Qy 1261 AGACCCAGAGAGGTGGCTTCGGACGAGCGCTTCGACAGAAATTCAGGTGCTGCA 1320
Db 421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440
Qy 1321 GTGCTGAAATTCGCGGCTTCTCAACAGAGATGTAGCCAGCAACTGTGTCGCCGT 1380
Db 441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460
Qy 1381 CAGTCTTTGAGCAGCGCATTTTGTGGATTGTCTCATGCGCTCATGCGCTCTGGA 1440
Db 461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480
Qy 1441 GCATCTCTGCTTAAATCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
Db 481 AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 500
Qy 1501 CGTGACCTGAGGTGCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1554
Db 501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518
RESULT 3
AAV22239
1D AAV22239 standard; Protein; 518 AA.
XX
AC AAV22239;
DE 20-SEP-1999 (first entry)
DE Human CSP56, aspartyl-type protease, protein sequence.
KW Metastatic marker protein; human; cancer metastasis; breast cancer;
KW colon cancer; diagnosis; therapy; tumour; metastatic potential;
KW CSP56; aspartyl-type protease.
XX Homo sapiens.
OS WO9934004-A2.
PN 08-JUL-1999.
PD 24-DEC-1998; 98WO-US27608.
PF 31-DEC-1997; 97US-0070112.
PR (CHIR) CHIRON CORP.
PA Giese K, Xin H;
FI WPI; 1999-430248/36.
DR N-PSDB; AAX84708.
XX
PT New polynucleotides associated with cancer metastasis
XX Claim 4; Page 78-80; 80pp; English.
CC This sequence represents a polypeptide of the invention, and is
CC an aspartyl-type protease, designated CSP56. The polynucleotides (PNS) of
CC the invention encode metastatic marker protein variants. The PNS and
CC polypeptides can be used as markers for cancer metastasis. The products
CC can be used for identifying metastatic tissue or metastatic potential of
CC a tissue, e.g. breast or colon tissue. They can also be used for
CC screening test compounds for the ability to suppress the metastatic
CC potential of a tumour. The products can be used for developing products
CC for the therapy of cancers, particularly breast or colon cancer.

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841


```

Db      301  ileValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320
QY      961  GAAGCTGTGGCCCGCATCTCTGATCCAGAAATCTCTGATGTTCTCTACTGCTGCTG 1020
Db      321  GluAlaValAlaArgAlaSerLeuLeuProGlnPheSerAspGlyPheThrGlySer 340
QY      1021  CAGCTGGCGTCTGGAGCAATTCGGAACACCTTGGTCTTACTTCCTAAATCTCCATC 1080
Db      341  GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrrPheProLysIleSerIle 360
QY      1081  TACCTGAGAGATCAGAACTCCAGAGAGTCATCCGTATACAACTCCGCTCAGCTTAC 1140
Db      361  TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTy 380
QY      1141  ATTCAAGCCCATGATGGGGCGCGCTCAATATCAATGTTACCAATTCGGCATTCGCCA 1200
Db      381  IleGlnPrometMetGlyAlaGlyLeuAsnTyroIuCysTyrrArgPheGlyIleSerPro 400
QY      1201  TCCACAAATGGCTGGTGTGATCGGTGCCAGCGTGAIGAGGGCTTCTACGTCACTTCGAC 1260
Db      401  SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrrValIlePheAsp 420
QY      1261  AGAGCCCAAGAGAGGGTGGCTTCCAGAGAGCCCTCTGCAGAAATTCAGGTGTGCA 1320
Db      421  ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440
QY      1321  GTGCTGTAATTCGGGGCTTTCTCAACAGAGATGATGACGCAACTCTGTGCCCGCT 1380
Db      441  ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460
QY      1381  CAGTCTTTGAGGAGCCCATTTTGTGATGTGTCCTATGCGTCAATGAGCGTCTGTGGA 1440
Db      461  GlnSerLeuSerGluProIleLeuTrpIleValSerTyrrAlaLeuMetSerValCysGly 480
QY      1441  GCATCTCTCTGTGTTAAATCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
Db      481  AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuLeuLeuProPheArgCysGlnArgPro 500
QY      1501  CGTGACCTGAGGTGCTCAATGATGAGTGCCTCTCTGCTGCTGCTGCTGCTGCTGCTG 1554
Db      501  ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 515

RESULT 7
ID      AA010628
AC      AA010628
XX      AA010628 standard; Protein: 518 AA.
XX      AA010628;
XX      AA010628;
DE      10-DEC-2001 (first entry)
DE      Human aspartyl protease 1 (hu-Aspl) protein.
KW      Human; aspartyl protease 1; Aspl; amyloid precursor protein; APP;
KW      Alzheimer's disease; AD; dementia; neurofibrillary tangle; gliosis;
KW      amyloid plaque; neuronal loss; proteolytic; neurotrophic; neuroprotective;
KW      chromosome 21.
OS      Homo sapiens.
FH      Key
FT      Peptide
FT      /label= Signal_peptide
FT      Protein
FT      /note= "Mature human aspartyl protease 1"
FT      Domain
FT      469..492
FT      /label= Transmembrane_domain
XX      GB2357767-A.
PN
XX
PD      04-JUL-2001.
XX
PF      22-SEP-2000; 2000GB-0023315.
XX

```

```

PR      23-SEP-1999; 99US-0155493.
PR      23-SEP-1999; 99US-0404133.
PR      23-SEP-1999; 99WO-US20881.
PR      13-OCT-1999; 99US-0416901.
PR      06-DEC-1999; 99US-0169232.
XX      (PHAA ) PHARMACIA & UPJOHN CO.
XX      Bienkowski MJ, Gurney M;
XX      WPI: 2001-444268/48.
XX      N-PSDB; AAD17864.
XX      Polypeptide comprising fragments of human aspartyl protease with.
XX      amyloid precursor protein processing activity and alpha-secretase
XX      activity, for identifying modulators useful in treating Alzheimer's
XX      disease.
XX      Claim 36; Fig 1; 187pp; English.
XX      The patent discloses human aspartyl protease 1 (hu-Aspl) or modified
XX      Aspl proteins which lack transmembrane domain or amino terminal
XX      domain or cytoplasmic domain and retains alpha-secretase activity
XX      and amyloid protein precursor (APP) processing activity. The proteins
XX      of the invention are useful for assaying hu-Aspl alpha-secretase
XX      activity, which in turn is useful for identifying modulators of
XX      hu-Aspl alpha-secretase activity, where modulators that increase
XX      hu-Aspl alpha-secretase activity are useful for treating Alzheimer's
XX      disease (AD) which causes progressive dementia with consequent
XX      formation of amyloid plaques, neurofibrillary tangles, gliosis and
XX      neuronal loss. Hu-Aspl protease substrate is useful for assaying
XX      hu-Aspl proteolytic activity, by contacting hu-Aspl protein with
XX      the substrate under acidic conditions and determining the level of
XX      hu-Aspl proteolytic activity. The present sequence is Aspl protein
XX      from human. Aspl gene is localised on chromosome 21.
XX      Sequence 518 AA:
XX      SQ
XX      Alignment Scores:
XX      Pred. No.: 4, 01e-213 Length: 518
XX      Score: 2687.00 Matches: 518
XX      Percent Similarity: 100.00% Conservative: 0
XX      Best Local Similarity: 100.00% Mismatches: 0
XX      Query Match: 80.47% Indels: 0
XX      DB: 22 Gaps: 0
XX      US-09-806-194-1 (1-1804) x AA010628 (1-518)
QY      1 ATGGCGCGACTGGCCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
Db      1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 20
QY      61 GCCCGGAGCTGGCCCGCCCGCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Db      21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40
QY      121 CGCGTAGTTGCGCCACCCCGGAGCCCGGAGCCCGCTGCGGAGCCCGAGCGGCTTG 180
Db      41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60
QY      181 GCGCTGCGCTGGAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db      61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80
QY      241 GTAGCAACCTCGAGGGGACTCTGCGCGCGCTACTACTCTGAGATGCTCATCGGACC 300
Db      81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrrLeuGluMetLeuIleGlyThr 100
QY      301 CCGCGCGAGAGCTACAGATCTCGTTGACACTGGAAGCAGTAACTTTGCGGTGGCAGGA 360
Db      101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120
QY      361 ACCCGCGACTCTCTACATAGACACAGTACTTTGACACAGAGAGGTCTAGCACATACCGTCC 420

```

DB 121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerThrTyrArgSer 140
QY 421 AAGGCTTTGACGTCACAGTCAAGTACACACAGGAGCTGCGGCTTCGTTGGGAA 450
DB 141 LysGlyPheAspValThrValLysTyrThrGlnGlySerThrGlyPheValGlyGlu 160
QY 481 GACCTCGTCACCATCCCAAGGCTTCAATACTCTTTCTTCTTCACAGATGCCACTAT 540
DB 161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 190
QY 541 TTTGATCAGAGATTTCTTTTCCTGGGATTAATGGAATGATCTGCGCTAGCT 580
DB 181 PheGluSerGluAsnPhePheLeuProGlyIleLysTyrAsnGlyIleLeuGlyLeuAla 200
QY 601 TATGCCACACTTGGCAGCCATCAAGTCTCTGAGACCTTCTTCCGACTCCCTGGTACA 560
DB 201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220
QY 661 CAACCAACATCCCAAGGTTTCTCCATGCAGATGCTGAGCGGCTGCGCGTGGCT 720
DB 221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240
QY 721 GGATCTGGGACCAACGGAGGTAGTCTTCTGCTGGTGGAAATGAACCAAGTTCTATAA 760
DB 241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyIleGluProSerLeuTyrLys 260
QY 781 GGAGATCTGGTATACCCCTATTAAAGAGAGTGGTACTACAGATAGAAATCTCAAA 840
DB 261 GlyAspIleThrTyrThrProIleLysGluGluThrTyrGlnIleGluIleLeuLys 280
QY 841 TTGGAAATTTGGAGCCCAAGGCTTAATCTGCACTGCAGAGATATAACGACAGACAGGC 900
DB 281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300
QY 901 ATCGTGGACAGTGCACACCGCTGCTGGCTGCCAGAGGCTTTGATGGGCTGCTG 960
DB 301 IleValAspSerGlyThrLeuLeuArgLeuProGlnLysValPheAspAlaVal 320
QY 961 CAAGCTGTGGCCCGGCACTCTGATCCAGAAATCTCTGATGTTCTGAGCTGATGCT 1020
DB 321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTyrThrArgLys 340
QY 1021 CAGCTGGCGTGTGACAGCAATTCGGAACACCTTGGTCTTACTTCCCTAAATCTCCAC 1080
DB 341 GlnLeuAlaCysTyrThrAsnSerGluThrProTyrPheProLysIleSerIle 1100
QY 1081 TACCTGAGAGATGAGAACTCCAGCAGGTTCATTCCGTATCAGAACTCGCTCAGCTTAC 1140
DB 361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380
QY 1141 ATTCAGCCATGATGGGGCGGCGCTGAATATGATGTTACCGATTCGGATTCGCCCA 1200
DB 381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400
QY 1201 TCCCAAAATGGCGTGTGATCGGTCGCCAGCGTGTGAGGCGCTCTACGTCATCTTCGAC 1260
DB 401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420
QY 1261 ACAGCCAGAGAGGTTGGCTTGGCGAGCGGCTGTCAGAAATTCAGGCTGCTGCA 1320
DB 421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAla 440
QY 1321 GTGTCTGAAATTCGGGCGCTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCGCGCT 1380
DB 441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460
QY 1381 CAGTCTTTGAGGAGCCCATTTTGTGATGTCCTATGCGCTCATGAGCGTCTGTGA 1440
DB 461 GlnSerLeuSerGluProIleLeuTyrIleValSerTyrAlaLeuMetSerValCysGly 480
QY 1441 GCATCTCTCTGCTTAATGCTCTCTGCTGCTGCGCTTCCGCTGAGCGTCCGCC 1500

DB 481 AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuLeuProPheArgCysGlnArgArgPro 500
QY 1501 CGTCACCTGAGGTCGTCAAATGATGAGTCCTCTCTGTCAGACATCGCTGGAAA 1554
DB 501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518
RESULT 8
AAE10656
ID AAE10656 standard; Protein; 518 AA.
XX AAE10656;
AC AAE10656;
XX 10-DEC-2001 (first entry)
DI Human-Asp 1 protein lacking TM domain and containing (His)6 tag.
DE Human; aspartyl protease 1; Aspl; amyloid precursor protein; APP;
KW Alzheimer's disease; AD; dementia; neurofibrillary tangle; gliosis;
KW amyloid plaque; neuronal loss; proteolytic; neurotrophic; neuroprotective.
XX Homo sapiens.
OS Synthetic.
XX GB2357767-A.
PN 04-JUL-2001.
PD 22-SEP-2000; 2000GB-0023315.
XX 23-SEP-1999; 99US-0155493.
PR 23-SEP-1999; 99US-0404133.
PR 23-SEP-1999; 99WO-US20881.
PR 13-OCT-1999; 99US-0416901.
PR 06-DEC-1999; 99US-0169232.
XX (PHAA) PHARMACIA & UPJOHN CO.
ZA Bienkowski MJ, Gurney M;
XX WP1; 2001-444208/48.
XX Polypeptide comprising fragments of human aspartyl protease with
PT amyloid precursor protein processing activity and alpha-secretase
PT activity, for identifying modulators useful in treating Alzheimer's
XX disease -
PS Example 14; Page 155-156; 187pp; English.
XX The patent discloses human aspartyl protease 1 (hu-Aspl) or modified
CC Aspl proteins which lack transmembrane domain or amino terminal
CC domain or cytoplasmic domain and retains alpha-secretase activity
CC and amyloid protein precursor (APP) processing activity. The proteins
CC of the invention are useful for assaying hu-Aspl alpha-secretase
CC activity, which in turn is useful for identifying modulators of
CC hu-Aspl alpha-secretase activity, where modulators that increase
CC hu-Aspl alpha-secretase activity are useful for treating Alzheimer's
CC disease (AD) which causes progressive dementia with consequent
CC formation of amyloid plaques, neurofibrillary tangles, gliosis and
CC neuronal loss. Hu-Aspl protease substrate is useful for assaying
CC hu-Aspl proteolytic activity, by contacting hu-Aspl protein with
CC the substrate under acidic conditions and determining the level of
CC hu-Aspl proteolytic activity. The present sequence is human Asp 1
CC protein lacking a transmembrane (TM) domain and containing (His)6
CC tag. This sequence is generated from human Asp 1 protein by the
CC deletion of its C-terminal TM domain and addition of hexa-histidine
CC tag at its C-terminus.
XX SQ Sequence 518 AA;
Alignment Scores: 4.01e-213 Length: 518
Pred. No.: 2687.00 Matches: 518
Score: 100.00% Conservative: 0
Percent Similarity:

QY	61	GC	CCCGGAGCTGGCCCGCCCGCCTTTCAGCTGC	CCCTCCCGSGTGGCCGCGCCACGAAC	120
DB	21	Ala	ProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn	40	
QY	121	CG	TACTTGGCGCCACCCCGGGAGCCCGGAGCCCTCGGAGCGCCAGCGCGGTG	180	
DB	41	Arg	ValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu	60	
QY	181	GC	GCTCCGCTGGAGCGTCGCTGGCGTCCCGCGCGCGCGCCCACTTCTTCGGCCATG	240	
DB	51	Ala	LeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet	80	
QY	241	GT	AGACAACCTGGAGGGGACTCTGC	CGCGCTACTACCTGGAGATGCTGATCGGAGCC	300
DB	81	Val	AspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr	100	
QY	301	CCCCG	CAGAAAGCTACAGATTCGCTGTGACACTGGAAGCAGTAACTTTGCGCTGGCAGGA	360	
DB	101	Pro	ProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly	120	
QY	361	AC	CGCGCACTCTACATAGACAGTACTTTGACAGAGAGAGTCTAGACAAIACCGGTCC	420	
DB	121	Thr	ProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer	140	
QY	421	AAG	GGCTTTCAGCTCAGTGAAGTACACACAAGAAAGCTGGACGGGCTCGTGGGGAA	480	
DB	141	Lys	GlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu	160	
QY	481	GAC	TGCTCACATFCCCAAGCGCTTCAATACTCTTTCTTGTCACATGGCCACATT	540	
DB	161	Asp	LeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle	180	
QY	541	TTT	GAATCAGAGAATTTCTTTTGTGCTGGGATTAAATGGAATGGAACTACTTGGCTAGCT	600	
DB	181	Phe	GluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla	200	
QY	601	TAT	GCCACACTTGCCACGCCCAAGTCTCTGAGAGACCTCTTCGACTCCCTGGTGACA	660	
DB	201	Tyr	AlaThrLeuAlaLysProSerSerLeuGluThrPhePheAspSerLeuValThr	220	

661	CAAGCAACATCCCCACAGCTTTCTCCATGCAGATGTGTGGAGCCGGCTIGCCCGTGTGCT	720
QY		
221	GlnAlaAsnIleP-oAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla	240
DB		
721	GGATCTGGGACCAACGAGGTAGCTTCTCTTGGGTGGAAITGAACCAAGTTGTGTATAAA	780
QY		
24:	GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuIleTyrLys	260
DB		
781	CGACACATCTGGTATACCCCTATTAGGAGAGAGTGGTACTACCACATAGAATTCGAAA	840
QY		
261	GlyAspIleThrPtyrThrProIleLysGluGluTyrTyrGlnIleGluLeuLeuLys	280
DB		
841	TTGCAATTCGAGGCCCAAGCCCTTATCTGGACTGCAGAGACTATAAGCGACACAGAAGCC	900
QY		
281	LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla	300
DB		

QY	961	AAGCTGTGGCCCGCATCTTGATTCGCAAAATTCTCTGATGGTTCTCGACTGGGTCC	1020
DG	301	IleValAspSerGlyThrLeuLeuAlaArgLeuProGlnLysValPheASPAlaVal	320
DB	321	GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer	340
QY	1021	CAGCTGGCGTGTGGACGAATTCGGAACACACTTGGTCTTACTTCCTAAAAATCTCCATC	1080

DO 341 GlnLeuAlaCysTrpPheAsnSerGluImpTolIpsSerIyrPheTrrGlysIleSerIle 360
 OY 1081 TACTGAGATGAGAACTCCACGAGTCATTCCGGTATCAACAATCCTCCCTACGTTTAC 1140
 |||||
 DB 361 TyrLeuArgaspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

QY 1141 ATTACGCCATGATGGCGCGGCTGAATATGAATGTTACCATTCGGCAATTCGCCA 1200
 DB 381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400
 QY 1201 TCCCAAAATGCGTGGTGAATCGGTGCCACGGTGGAGGCTTCTACGTCATCTCCAC 1200
 DB 403 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420
 QY 1261 AGACCCAGAAAGAGGTGGGTTCGACGAGCCGCTGTCAGAAATGAGGTGCGCA 1320
 DB 421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440
 QY 1321 GTGCTGAAATTCGGGCCCTTCTCAACAGAGATGAGCCACACATGTCGCCGCT 1380
 DB 441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460
 QY 1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTCTGTCATGCGCTGAGCGCTG 1440
 DB 461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480
 QY 1441 GCATCTCTCTGTGTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
 DB 481 AlaIleLeuLeuValIleValIleValIleValIleValIleValIleValIleVal 500
 QY 1501 CGTGACCTGAGGTGCTGCAATGATGAGTCCTCTCTGCTGACACATCGCGGAA 1554
 DB 501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgIleProlys 518
 RESULT 10
 AAE06858
 ID AAE06858 standard; Protein: 518 AA.
 XX
 AC AAE06858;
 CC
 DT 23-OCT-2001 (first entry)
 XX
 DE Human aspartyl protease 1 (Hu-Asp1) protein.
 KW Human; aspartyl protease 1; Asp 1; beta-amyloid precursor protein; APP;
 KW beta-secretase; Alzheimer's disease; dementia; amyloid plaque; gliosis;
 KW neurofibrillary tangle; neuronal loss; amyloid-beta peptide; nontropic;
 KW neuroprotective; antisense therapy; gene therapy; chromosome 21.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..20
 FT Protein /label= Signal_peptide
 FT Domain /note= "Mature human aspartyl protease 1 (Hu-Asp1)"
 FT /label= Transmembrane_domain
 XX
 PN W020C150829-A2.
 XX
 PD 19-JUL-2001.
 XX
 PF 09-MAY-2001; 2001WO-IB00799.
 XX
 PR 09-MAY-2001; 2001WO-IB00799.
 XX
 PA (BIEN/) BIENKOWSKI M J.
 PA (GURN/) GURNEY M E.
 PA (HEIN/) HEINRIKSON R L.
 PA (PARO/) PARODI L A.
 PA (YANR/) YAN R.
 XX
 PI Bienkowski MJ, Gurney ME, Heinrikson RL, Parodi LA, Yan R:
 XX
 DR WPI: 2001-483072/52.
 DR N-PSDB: AAD13020.
 XX
 PT Novel purified polypeptide comprising fragment of mammalian aspartyl

PT protease 2, lacking Asp2 transmembrane domain and retaining beta
 PT secretase activity of Asp2 useful for identifying inhibitors of Asp2
 PT activity
 XX
 PS Example 2; Fig 1; 185pp; English.
 XX
 CC The invention relates to human aspartyl proteases (Hu-Asp), beta-amyloid
 CC precursor protein (APP) isoforms and their corresponding DNA molecules.
 CC Human aspartyl proteases can act as beta-secretase proteases useful for
 CC treating Alzheimer's disease. APP isoforms are useful for identifying
 CC modulators of amyloid-beta peptide production, for use in designing
 CC therapeutics for the treatment and prevention of Alzheimer's disease,
 CC dementia, formation of amyloid plaques, neurofibrillary tangles, gliosis
 CC and neuronal loss. APP isoforms are also used in methods for identifying
 CC inhibitors and modulators of human Asp2 activity. The invention relates
 CC to a method for identifying agents that modulate the activity of human
 CC aspartyl protease Asp2. Amyloid-beta peptides obtained from APP are used
 CC as a means to screen in cellular assays for the inhibitors of beta- and
 CC gamma- secretase. Hu-Asp DNA fragments are useful as probes or primers in
 CC polymerase chain reactions (PCR). The probes are useful for detecting
 CC Hu-Asp nucleic acids in in vitro assays and in Northern and Southern
 CC blots. The present sequence is human aspartyl protease 1 (Hu-Asp1).
 CC Hu-Asp 1 gene is localised on chromosome 21.
 XX
 SQ Sequence 518 AA;
 Alignment Scores:
 Prod. No.: 4.01e-213 Length: 518
 Score: 2687.00 Matches: 518
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 80.47% Indels: 0
 DB: 22 Gaps: 0
 US-09-806-194-1 (1-1804) x AAE06858 (1-518)
 QY 1 ATGGGGCCACTGGCGCGGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
 DB 1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20
 QY 61 GCGCCGAGCTGGCGCGCGCGCTTCAGCTGCGCCCTCCGGTGGCGCGCGCGCGCGCGAAC 120
 DB 21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40
 QY 121 CGCTAGTTGCGCCGCCACCGCGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 180
 DB 41 ArgValAlaAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAlaGlyLeu 60
 QY 181 GCGCTCGCCCTGGAGCTTCCCTGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
 DB 61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80
 QY 241 GTAGACAACCTGCAGGGGACCTGCGCGCGCGCTACTACCTGGAGATCGTGAATCGGGACC 300
 DB 81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100
 QY 301 CCCCCCAGAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACITTTGCCGTGCGAGGA 360
 DB 101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120
 QY 361 ACCCGCAGCTCCTACATAGACAGCTACTTTGACAGAGAGGCTTAGCACATACCGCTCC 420
 DB 121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140
 QY 421 AAGGGCTTTGAGCTCAGCTGAGTACACAGAGGAGCTGCGCGCGCTGCTGCTGGGAA 480
 DB 141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTyrThrGlyPheValGlyGlu 160
 QY 481 GAGCTCGTCACCATCCCCAAAGGCTTCAATACCTTCTTCTTCTTCTTCTTCTTCTTCTT 540
 DB 161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180
 QY 541 TTGGAATCAGAGAAATTTCTTTTGCTGGGATTAAATGAATGAATGAATGAATGAATG 600

```

Db      181 PheLusSerGluAsnPhleLeuProGlyLeLysrAsnGlyLeLeGlyLeuAla 300
Qy      601 TATGCCACATTCGCCAGCCATCAAGTCTCTCGAGACCCITCTCGACTCCCTGGTGACA 660
Db      201 TyrAlaThrLeuAlaLysProSerSerLeuGluThrPhePheAspSerLeuValThr 220
Qy      661 CAAGCAACATCCCAACGTTTCTCCATCCAGATGTGTGGAGCCCGCTGGCGTTCGT 720
Db      221 GlnAlaAsnLeProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240
Qy      721 GCATCTGGGACCAACGAGGAGTCTTCTCTGGTGGATTGAACCAAGTTGTATATAA 780
Db      241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyLeuProSerLeuThrLys 260
Qy      781 GGAGACATCTGGTATACCCCTATTAAGGAAGATGTGTACTACAGATAGAATCTGANA 840
Db      261 GlyAspIleTrpThrProIleLysGluGluTrpTyrTyrGlnIleGluLeuLys 280
Qy      841 TTGGAATTTGAGGCCAAAGCCCTTAATCTGAGTGCAGAGATATAACGACAGAGCC 900
Db      281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluIyrAsnAlaAspLysAla 300
Qy      901 ATCGTGACAGTGGCACCCAGCCCTGCTGCGCTCCGCCAGAGGTGTGTTCATCGGTGT 960
Db      301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaVal 320
Qy      961 GAAGCTGTGCGCCGCGATCTCTGATTCAGAAATCTCTGATGTTCCTCGACTGGTCC 1020
Db      321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340
Qy      1021 CAGCTGCGCTGCTGGAGCAATTCGGAACACCTTGTCTTACTTCCTAAATCTCATC 1080
Db      341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360
Qy      1081 TACCTGACAGATGAGAACTCCAGCAGGTCTCTCGTATACAAATCTCGCTACCTTAC 1140
Db      361 TyrLeuArgAspGluAsnSerArgSerPheArgIleThrIleLeuProGlnLeuThr 380
Qy      1141 ATTCAAGCCCATGATGGGGCGCGCTGAATATGAATGTACCGATTCGGCATTCGCCA 1200
Db      381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400
Qy      1201 TCCACAAATTCGGTGTGATCGGTGGCCAGCGTGTACAGGCGCTTACATCTTCAGAC 1260
Db      401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420
Qy      1261 AGACCCAGAGAGGGTGGCTTCGACGCGAGCCCTGTCCAGAAATTCAGTGTGTGCA 1320
Db      421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440
Qy      1321 GTGTCTGAATTCGGGCGCTTCTCAACAGAGATGTACGACAGCACTGTGTCCCGCT 1380
Db      441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnGlyValProAla 460
Qy      1381 CAGTCTTTGACGAGCCCATTTGTGGATTGTGCTCTATCGCTCATGAGCGTCTGGAA 1440
Db      461 GlnSerLeuSerGluProIleLeuThrPheValaSerTyrAlaLeuMetSerValCysGly 480
Qy      1441 GCCATCTCTGTTGTTAAATCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
Db      481 AlaIleLeuLeuValLeuIleValLeuLeuLeuProPheArgCysGlnArgArgPro 500
Qy      1501 CGTGACCTGAGTGCATCATGATGATGATGATGATGATGATGATGATGATGATGATG 1554
Db      501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518

```

RESULT 11

AAU06602

ID AAU06602 standard; Protein: 518 AA.

XX

AC AAU06602;

XX

```

24-OCT-2001 (first entry)
Human Aspartyl protease 1 (Asp1).
Human: Aspartyl protease; Asp1; Asp2; beta-secretase; nototropic;
neuroprotective; amyloid protein precursor; APP; Alzheimer's disease;
amyloid-beta; Abeta.
XX Homo sapiens.
OS WO200149098-A2.
PN 12-JUL-2001.
XX 09-MAY-2001; 2001WO-IB00798.
XX 09-MAY-2001; 2001WO-IB00798.
XX (BIEN/) BIENKOWSKI M J.
PA (GUEN/) GURNEY M E.
PA (HEIN/) HEINRIKSON R L.
PA (PARC/) PARODI L A.
PA (YANR/) YAN R.
XX Bienkowski MJ, Gurney ME, Heinrikson RL, Parodi LA, Yan R;
WPI: 2001-502549/55.
DR N-PSDB; AAS11516.
XX Novel purified polypeptide comprising fragment of mammalian aspartyl
protease 2, lacking Asp2 transmembrane domain and retaining beta
secretase activity of Asp2 useful for identifying inhibitors of Asp2
activity -
XX Example 2; Fig 1; 185pp; English.
XX The invention relates to a purified polypeptide comprising a fragment of
mammalian aspartyl protease (Asp2) protein which lacks the Asp2
transmembrane domain and the Asp2 protein, and where the polypeptide and
the fragment retain the beta-secretase activity of the mammalian Asp2
protein. The invention also details polynucleotides for the Asp
proteins and vectors expressing them, and a polypeptide (isoform of
amyloid protein precursor (APP)) comprising the amino acid sequence of an
APP or its fragment containing an APP cleavage site recognizable by a
mammalian beta-secretase, and further comprising two lysine residues at
the carboxyl terminus of the amino acid sequence of the mammalian APP or
APP fragment. Also included in the invention are methods of identifying
modulators or inhibitors of Asp2. Modulators and inhibitors of Asp2 are
useful for treating Alzheimer's disease. APP is useful in methods for
identifying inhibitors or modulators of human Asp2 activity and
amyloid-beta (Abeta) peptide production. APP is also useful in designing
therapeutics for the treatment or prevention of Alzheimer's disease.
APP comprising the APP-Sw-beta-secretase peptide sequence (NLDA), which
is associated with increased levels of Abeta processing is useful in
assays relating the Alzheimer's research. The expression vector is useful
for recombinantly expressing APP. Nucleic acids that hybridize to
APP oligonucleotides are useful as probes or primers. The probes are
useful for detecting Hu-APP nucleic acids in in vitro assays and in
Northern and Southern blots. The present sequence is human Asp1.
Sequence 518 AA:
Alignment Scores:
Pred. No.: 4,01e-213 Length: 518
Score: 2687.00 Matches: 518
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 80.47% Indels: 0
DB: 22 Gaps: 0
US-09-806-194-1 (1-1804) x AAU06602 (1-518)
Qy 1 ATGGGCGCACTGGCCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGCC 60

```

Db 1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20
Qy 61 GCGCCGAGCTGGCCCGCCGCTTACACGTCGCCCTCGGGTGGCGCCGCGCCAGAAC 120
Db 21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40
Qy 121 CGGTAGCTTGGCCGACCCCGGAGCCGCGGACCCCTCGCGAGCGCCAGCCGACGCTTCG 180
Db 41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLe 60
Qy 181 GCGTCGCCCTGGAGCTGCGCTGGCTGCCCGCGCGCGCCGCCACACCTCTTGGCCGATG 240
Db 61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGluAlaAlaAsnPheLeuAlaMet 80
Qy 241 GTAGACACCTGCGAGCGGAGCTTGGCCGCGGCTACTACCTGGAGATGCTGATCGGACT 300
Db 81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuTlleGlyThr 320
Qy 301 CCCCCGAGAGCTACAGATCTCGTTGACACTGGAGCAGTAACCTTGGCGTGGTAGGA 360
Db 101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 420
Qy 361 ACCCGCAGCTCTACATAGACAGCTACTTGGACACAGAGGCTGTACACATACCGCTTC 420
Db 121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerThrTyrArgSer 440
Qy 421 AAGGGCTTGGCTCAGCTAGTAGTACACACAGGAGCTGGAGCGGCTGCTGGGGA 480
Db 141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTyrThrGlyPheValGlyGlu 500
Qy 481 GACCTGCTCACCATCCCAAGGCTTCAATACTTCTTCTTCTGACATTCGCACATTC 540
Db 161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 560
Qy 541 TTGGAATCAGAGAATTTCTTTTTCCTGGGATTAATGGAATGGAATGGAATGGAATGGA 600
Db 181 PheGluSerGluAsnPhePheLeuProGlyIleLysTyrAsnGlyIleLeuGlyLeuAla 620
Qy 601 TATGCCACACTGCCAGCATCAAGTTCTCTGGAGACCTTCTTCGACCTCCCTGGTACA 660
Db 201 TyrAlaThrLeuAlaLysProSerSerLeuGluThrPhePheAspSerLeuValThr 720
Qy 661 CAAGCAACATCCCAACGTTTCTCCATGATGTGGAGCGCGCTGCGCGGTTGCT 720
Db 221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 740
Qy 721 GGATCTGGGACCAACGGAGTAGTCTTGTCTGGTGGAAATGAACCAAGTTCTGTATAA 760
Db 241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 780
Qy 781 GGACACATCTGGTATACCCCTATTAAAGGAAGTGGTACTACAGATAGAAATCTGTAAA 840
Db 261 GlyAspIleThrTyrThrProIleLysGluGluTyrTyrGlnIleGluIleLeuLys 860
Qy 841 TTGGAATGGAGCCCAAGCTTAATCTGAGCTGAGAGATATAACGAGACAGAGCC 900
Db 281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 920
Qy 901 ATCGTGGACAGTGGACACGCTGCTGGCTGCCCGAGAGGTGGTGGTGGTGGTGGTGG 960
Db 301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 980
Qy 961 GAAGCTGTGGCCCGGCACTCTCTGATTCAGAAATCTCTGATGGTTTCTGACCTGGTCC 1020
Db 321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheThrTrpTrpGlySer 1040
Qy 1021 CAGCTGGCGTGGAGCAATTCGGAACACCTTGGCTTACTTCCTCAAAATCTCCATC 1080
Db 341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 1100
Qy 1081 TACCTGAGATGAGAACTCCAGCAGCTCATTCGGTATCAATCTCGCTCAGCTTAC 1140

Db 361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380
Qy 1141 ATTCAGCCCATGATGGGCGCGGCTGAATATTAATGATTAATGATTAATGATTAATGATTA 1200
Db 381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400
Qy 1201 TCCACAAATGGCTGGTATCGGTCCACGGTGTAGAGGGCTTCTAGCTCATCTTCGAC 1260
Db 401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420
Qy 1261 ACAGCCCAAGAGAGGCTGGCTTCGACGAGCCCTGTGCAGAAATTCAGAGGTGCTGCA 1320
Db 421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440
Qy 1321 GTGCTGAAATTCGCGGCCCTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCGCCGCT 1380
Db 441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460
Qy 1381 CAGTCTTTGAGCGACCCCATTTTGGGATTTGTCTATGCTATGCTATGCTATGCTATGCT 1440
Db 461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480
Qy 1441 GCCATCTCTCTGCTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
Db 481 AlaIleLeuLeuValIleValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 500
Qy 1501 CQTGACCTGAGGTCGTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1554
Db 501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518

RESULT 12
AAU07201
ID AAU07201 standard; Protein; 518 AA.
XX
AC AAU07201:
D7
D7 24-OCT-2001 (first entry)
XX
DE Human aspartyl protease 1 (Asp-1).
XX
KW Human; aspartyl protease 1; Asp-1; nootropic; neuroprotective;
KW aspartyl protease 2; Asp2; amyloid protein precursor; APP;
KW beta-secretase; Alzheimer's disease.
XX
OS Homo sapiens.
XX
PN W0200149097-A2.
XX
PD 12-JUL-2001.
XX
PF 09-MAY-2001; 2001WO-1B00797.
XX
PR 09-MAY-2001; 2001WO-1B00797.
XX
PA (RIEN/) BIENKOWSKI M J.
PA (GURN/) GURNEY M E.
PA (HEIN/) HEINRIKSON R L.
PA (PARO/) PARODI L A.
PA (YANR/) YAN R.
XX
PI Bienkowski MJ, Gurney ME, Heinrikson RL, Parodi LA, Yan R;
XX
DR WPI; 2001-502548/55.
XX
DR N-PSDB: AAS11701.
XX
XX
PT Novel purified polypeptide comprising fragment of mammalian aspartyl
PT protease 2, lacking Asp2 transmembrane domain and retaining beta
PT secretase activity of Asp2 useful for identifying inhibitors of Asp2
PT activity
XX
XX Example 2; Fig 1; 185pp; English.
PS
XX The invention relates to a novel purified polypeptide comprising a

601	QY	TATGCCACACTTGCCCAAGCCATCAAGATTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA	660
201	Db	TyrAlaThrLeuAlaLysProSerSerLeuGluThrPhePheAspSerLeuValThr	220
661	QY	CAAGCAAAACATCCCCAACGTTTTCTCCATGCGACATGTGGAGCCCGCTGCCCGCTGCT	720
221	Db	GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla	240
721	QY	GGATCTGGGACCAACGGAGGTAGTCTTGTCITGGGTGGAAATCAACCAAGATTCTATAA	780
241	Db	GlySerGlyThrAsnGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys	260
781	QY	GSAGACATCTGATACCCCTATTAAAGGAAGACTGGTACTACCAGATAGAAATCTCGAAA	840
261	Db	GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys	280
841	QY	TTGGAAATTGGAGGCCAAAGCCCTTAATCTGGACTGCAGAGAGTATACCGAGACACAGGCC	900
281	Db	LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla	300

QY 501 ATCTGGACAGTGGCACGACGCTGGCTGGCCGACGAAAGGTGTATGATGCGGTGGTG 960

301 IEVATASPSEIGYIMI IMLLEUENKIGLEUFI OGIMLYSVAIFHEKSPATDVAIAT 320

961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAAATCTCTGAIGTTTCTGGACTGGGTCC 1020
|||||
321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340
|||||
1021 CAGCTGGGTGTGTGGACGAATTCGGAACACCTTGGTCTTACTTCCTCAAAATCTCCATC 1080

DE 341 619649acystrprmtasuseetomumfrompseryytnefforysreseeire 300

QY	1081	IACCTGAGAGATGAGAACATCCACGAGGUAATCCGATATCAACAATCTCCCTCAGCTTAC	1141
Db	361	TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr	380
QY	1141	ATTCAGCCCATGATGGGGCCGCGCTGAATTATGAATGTTACGGATTTCGGCATTTCCCCA	1200
Db	381	IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro	400
QY	1201	TCACAAATTCGGCTGGTGTATCGGTGCACCGGTGATGAGGGCTTCTACGTCTCTTCGAC	1260
Db	401	SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp	420
QY	1261	AGAGCCCAAGAGGGTGGCTTCGAGCGAGGCCCTGTGTCACAAATTCGAGGTGCTGCA	1320
Db	421	ArgAlaGlnLysArgValGlyPheAlaIleAsnProCysAlaGluIleAlaGlyAlaAla	440
QY	1321	GTGCTGAAATTTCCGGGCTTCTCAACAGAGAGATGTAGCCAGCAACTGTGCCCGCGCT	1380
Db	441	ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla	460
QY	1361	CAGTCTTTGAGCGAGCCCAATTTGTGGATTGTCTCTATCGCGCTCATGAGCGTCTGTGA	1440
Db	461	GlnSerLeuSerGluProIleLeuTrpIleVal-SerTyrAlaLeuMetSerValCysGly	480
QY	1441	GCCATCCTCTTCTTAATCGTCTCTGCTGCTGCTGCTGCCGTTCCGGTGTACAGCTCGCCC	1500
Db	481	AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuProPheArgCysGlnArgArgPro	500
QY	1501	CGTGACCTTGAGTCTCGTCAATGATGAGTCTCTCTGTGTGTCAGACATCGCTGGA	1554
Db	501	ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys	518
RESULT 13			
AAE02580			
1D ID AAE02580 standard; Protein; 518 AA.			
XX AC AAE02580;			
XX AC AAE02580;			
XX DT 10-AUG-2001 (first entry)			
XX			

DE Human aspartyl protease 1 (Asp 1).

QY 1081 TACCTGACGATGAGACATCCAGCAGGTCTATCCGATACAACTCCGCCACGCTTAC 1140
 Db 361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrLeuPheGlnLeuTyr 380
 QY 1141 ATTCAGCCCATGATGGGCGGCGCTGAATTAATGATGTTACCGATTCGGCATTCGCA 1200
 Db 381 IleGlnPrometMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400
 QY 1201 TCCCAAAATGGCTGGTGAATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1260
 Db 401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420
 QY 1261 ACAGCCCAAGAGGCTGGCTCCACGAGGCGCTCTGCAGAAATTCGAGTGGTGGCA 1320
 Db 421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaGala 440
 QY 1321 GTGCTGAAATTCGGGCGCTTCTCAACAGAGGATGTAGCCCAACTGTGTCCCGGCT 1380
 Db 441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460
 QY 1381 CAGTCTTTCAGGAGCCCATTTTGTGGATGTCCTATGCGCTCATGAGGCTCTGTGGA 1440
 Db 461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480
 QY 1441 GCATCCCTCTTGTCTTAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
 Db 481 AlaIleLeuValLeuIleValLeuLeuLeuLeuLeuProPheArgCysGlnArgPro 500
 QY 1501 CSTGACCTGAGTGTCTCAATCATGATGATGATGATGATGATGATGATGATGATGAT 1560
 Db 501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgIlePlys 518
 RESULT 14
 AAE02608
 ID AAE02608 standard; Protein: 518 AA.
 AC AAE02608;
 DT 10-AUG-2001 (first entry)
 DE Human Aspartyl protease-1 (Asp-1) deltaTM (His)6 protein.
 KW Human: alpha-secretase; amyloid precursor protein; APP; therapy;
 KW Alzheimer's disease; anialzheimer's; aspartyl protease 1; Asp-1;
 KW beta-secretase; Asp-1 deltaTM (His)6 protein.
 CS Homo sapiens.
 CS Synthetic.
 XX WO200123533-A2.
 XX 05-APR-2001.
 XX 22-SEP-2000; 2000MO-US26080.
 XX 23-SEP-1999; 99US-0155493.
 XX 23-SEP-1999; 99MO-US20881.
 XX 13-OCT-1999; 99US-0418901.
 XX 06-DEC-1999; 99US-0169232.
 XX (PHAA) PHARMACIA & UPJOHN CO.
 XX Gurney M, Bienkowski MJ;
 XX WPI: 2001-290516/30.
 XX Enzymes that cleave the alpha-secretase site of the amyloid precursor
 XX protein, useful for the treatment of Alzheimer's disease -
 PS Example 14; Page 183-184; 189pp; English.
 XX The present invention relates to enzymes for cleaving the alpha-
 XX secretase site of the amyloid precursor protein (APP) and methods of

CC identifying those enzymes. The methods may be used to identify enzymes
 CC that may be used to cleave the alpha-secretase cleavage site of the APP
 CC protein. The enzymes may be used to treat or modulate the progress of
 CC Alzheimer's disease. The present sequence is human Aspartyl protease-1
 CC (Asp-1) deltaTM (His)6 protein which is used for the expression of
 CC pre-pro-human Aspartyl protease 1 (Asp1). This protein is obtained by
 CC replacing C-terminal transmembrane and cytoplasmic domains with a
 CC hexahistidine purification tag in the human Aspartyl protease 1.
 XX
 SQ Sequence 518 AA:
 Alignment Scores:
 Pred. No.: 4, 01e-213 Length: 518
 Score: 2687.00 Matches: 518
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 80.47% Indels: 0
 DB: 22 Gaps: 0
 US-09-806-194-1 (1-1804) x AAE02608 (1-518)
 QY 1 ARGGGCGACTGGCCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
 Db 1 MetGlyValAlaLeuAlaArgAlaLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20
 QY 61 GCCCGGAGCTGGCCCGCGCCCTTCAGCTGCCCTCCGGGTGGCGCGCCGACCAAC 120
 Db 21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40
 QY 121 CGGTAGTTGGCCCGACCCGGGACCGGACCCCTCGCGAGCGCCGACCGCGAGCGGTG 180
 Db 41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60
 QY 181 GCGCTCGCCCTGGAGCTCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCTGCGCTG 240
 Db 61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80
 QY 241 GTAGCAACTGCAGGGGACTCTGCGCGCGCTACTACTCTGGAGATGCTGATCGGGACC 300
 Db 81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100
 QY 301 CCCCCAGAGCTACAGATTCTCGTTGACATGGAAGCAGTAACTTTGCCGTGGCAGGA 360
 Db 101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120
 QY 361 ACCCGCACTCTACATAGACACAGTACTTTGACAGAGAGAGTCTAGACATACCGTCC 420
 Db 121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140
 QY 421 AAGGGCTTTGAGTGCAGCTGAGTACACAGAGAGCTGGACGGCTCTGTGGGAA 480
 Db 141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTyrThrGlyPheValGlyGlu 160
 QY 481 GACCTGCTCACCATCCCAAGGCTTCAATACCTTCTTGTGCAACATGCCACTATT 540
 Db 161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180
 QY 541 TTGTAATCAGAAATTTCTTTTGGCTGGGATTAATGGAATGGAATGGAATGGAATG 600
 Db 181 PheGluSerGluAsnPhePheLeuProGlyIleLysTyrPheAsnGlyIleLeuGlyLeuAla 200
 QY 601 TATGCCACACTTGGCCAAAGCATCAAGTCTCTGGAGACCTCTTCCACTCCCTGGTGACA 660
 Db 201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220
 QY 661 CAAGCAAAATCCCCAACGTTTCTCCATGAGATGTGTGGAGCGCGTTCGCCGTTCCT 720
 Db 221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240
 QY 721 GGATCTGGGACCAACCGAGGTAGTCTTCTCTGGTGGTGAATTAACCAAGTTTCTATATA 780
 Db 241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyIleGluProSerLeuTyrLys 260


```
Db 81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGlnMetLeuIleGlyThr 100
QY 301 CCCCCCAGAGCTACAGATCTTCGTTGACACTGGAAGCACTAACTTCGCCGTGCAGGA 360
Db 101 ProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120
QY 361 ACCCGGCACCTCTACATAGACAGCTACTTTGACACAGAGAGGTCTAGCACATACAGCTCC 420
Db 121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerThrTyrArgSer 140
QY 421 AAGSGCTTCACGCTCACAGTGAAGTACACACAGAGAGCTGGACGGCTTCGCTGGGGA 480
Db 141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpTrpGlyPheValGlyGlu 160
QY 481 GACCTCGTCACCATCCCAAGCGCTCAATACTCTTTCTTTCACAACTGCCACIAT 540
Db 161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180
QY 541 TTGCAATCAGAGAAATTCCTTTTGGCTGGGATTAATGAATGAATGATCTGGCGTAGCT 600
Db 181 PheIuSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200
QY 601 TATGCCACACTTCCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
Db 201 TyrAlaThrLeuAlaLysProSerSerLeuGluThrPhePheAspSerLeuValThr 220
QY 661 CAAGCAACATCCCAAGCTTTCTCCATGCAGATGTGTGGAGCGGCTTGCCCTTGCT 720
Db 221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240
QY 721 GGATCTGGGACCAACGAGGTAGTCTGTGCTGGTGGATTAACCAAGTTTGATAAA 780
Db 241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrIlys 260
QY 781 GGACATCTGGTATACCCCTATTAGGAGAGTGGTACTACCATATAAATTCIGAAA 840
Db 261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTrpGlnIleGluIleLeuLys 280
QY 841 TTGGAATTTGAGGCGCAAGCCCTTAATCTGGACTGCAGAGAGTAAACGACAGCAAGGCC 900
Db 281 LeuGluIleGlyGlyGlnSorLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300
QY 901 ATCGTGCACAGTGCACACAGCTGCTGCCCTGCCCGCCAGAGAGTGTTCATCGGAGTGS 960
Db 301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320
QY 961 GAAGCTGTGCCCGCGCATCTCTGATTCAGAGATTCCTGTATGGTTCCTGGACTGGGTTC 1020
Db 321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340
QY 1021 CAGCTGGCGTGTGGAGCAATTCGGAACACCTTGCTCTACTTCCTAAATCTCCATC 1080
Db 341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360
QY 1081 TACCTGAGAGATGAGAACTCCAGCAGCTCATTCGGTATCACAAATCGCTCAGCTTTAC 1140
Db 361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380
QY 1141 ATTCAGCCCATGATGGGGCGCCGCTGAATATGAATGTACCGATTCGGCAATTCCCCCA 1200
Db 381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400
QY 1201 TCCCAAAATCGCTGGTGTATCGGTGCACAGTGTACGGGTTCCTACGTCATCTTCGAC 1260
Db 401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420
QY 1261 AGAGCCAGAGAGGGTGGGCTTCGCGAGCGAGCCCTGTCCGAAAATTCAGGTGTCTCCA 1320
Db 421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440
QY 1321 GTGTCTGAAATTTCCGGGCGCTTCTTCAACAGAGGATGTAGCCAGCAACTGTGTCCCGCT 1380
```

```
Db 441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460
QY 1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTTGTCTATGCGCTCATGAGGCTCTGTGGA 1440
Db 461 GluSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480
QY 1441 GCCATCCTCCTTGTCTTAATCGTCTCTGCTGCTGCCGTTCGGTGTCCAGGTGCGCCC 1500
Db 481 AlaIleLeuLeuValLeuIleValLeuLeuLeuProPheArgCysGlnArgArgPro 500
QY 1501 CGTGACCTCGAGGTCGCTCAATGATGATGATGCTCTCTGGTCAGACATCGCTGGAAA 1554
Db 501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518
```

Search completed: September 25, 2003, 11:00:02
Job time : 146.5 secs